Department of Microbiology, Tumor and Cell Biology
Karolinska Institutet, Stockholm, Sweden

EPIGENETIC REGULATION OF HIGHER ORDER CHROMATIN CONFORMATIONS AND NETWORKS

Chengxi Shi

Karolinska Institutet

Stockholm 2013

Published by Karolinska Institutet. Printed by Larserics Digital Print AB

ISBN 978-91-7549-187-5

To my family

致我的家人

"All our knowledge begins with the senses, proceeds then to the understanding, and ends with reason. There is nothing higher than reason."

- Immanuel Kant
Critique of Pure Reason (1781)

ABSTRACT

Higher order chromatin conformations result from the packaging of the genome into the physical confines of the cell nucleus. Structural hallmarks of the nucleus influence the spatio-temporal behavior of genome underlying the regulation of genomic functions. Moreover, accumulated data show that the physical proximities between interphase chromatin fibers significantly contribute to the regulation of genomic transcription, replication and repair. The dynamic patterns of spatial crosstalk between genomic regions are, moreover, controlled by environmental cues to fine-tune gene transcription.

The studies in this thesis focus on the nature of higher order chromatin conformations and networks and their developmental regulation in mouse and human model systems. The thesis also includes the description of a novel technique that enables the visualization of higher order chromatin proximities in single cells at a resolution far beyond that of the microscope.

Specifically, we identified developmentally regulated genome-wide chromosomal interactomes impinging on the *H19* imprinting control region (ICR) in embryonic stem (ES) cells and derived embryoid bodies (EBs). The chromosomal interactomes appear poorly conserved between mouse and human. We further constructed chromosomal interaction networks with crosswise interacting pattern and present the modular topology of the human networks. The molecular glue connecting chromosomes to each other was identified as poly(ADP-ribose). TGFβ signaling was shown to rapidly rewire the chromosomal interaction networks by targeting a CTCF-PARP1 feed-back loop to decrease poly(ADP-ribose) levels in the nucleus.

We further captured a developmentally conserved imprinted interaction network, which is dependent on CTCF binding sites on the maternal *H19* ICR allele. This network was shown to function as a vehicle to transfer epigenetic states from *H19*/*Igf2* domain to other imprinted domains it interacts with. We propose the principle of non-allelic transvection of epigenetic states as a notable functional outcome of the physical contacts between chromatin fibers.

Finally, we invented Chromatin *In Situ* Proximity (ChrISP), which is a novel technique to identify and visualize proximities between chromatin fibers or between chromatin fiber as well as structural hallmarks in single cells at a high resolution. By employing the ChrISP technique we demonstrated that modification of epigenetic marks by environmental cues triggers large-scale changes in chromosome conformations.

It is concluded that higher order chromatin conformations and networks are epigenetically regulated by environmental cues and significantly contribute to the regulation of genomic functions during developmental and pathological processes.

LIST OF PUBLICATIONS

This thesis is based on the following original papers:

I. Kuljeet Singh Sandhu*, **Chengxi Shi***, Mikael Sjölinder*, Zhihu Zhao*, Anita Göndör, Liang Liu, Vijay K Tiwari, Sylvain Guibert, Lina Emilsson, Marta P Imreh, and Rolf Ohlsson (2009). Nonallelic transvection of multiple imprinted loci is organized by the *H19* imprinting control region during germline development.
Genes & Development, *23(22)*, 2598-2603.

II. Anita Göndör, Alejandro Fernandez Woodbridge, **Chengxi Shi**, Erik Aurell, Marta P Imreh, and Rolf Ohlsson (2010). Window into the complexities of chromosome interactome.
Cold Spring Harbor Symposia on Quantitative Biology, *75*, 493-500.

III. Anita Göndör, Alejandro Fernandez Woodbridge, Noriyuki Sumida, Xingqi Chen, Samer Yammine, Moumita Biswas, Olga Loseva, **Chengxi Shi**, György Stuber, Balazs Nemeti, Maria Israelsson, Thomas Helleday, Marta P Imreh, and Rolf Ohlsson. TGFβ regulates physical interactions between chromosomes mediated by poly(ADP-ribose).
Manuscript.

IV. Xingqi Chen*, **Chengxi Shi***, Samer Yammine*, Anita Göndör, Mariliis Tark-Dame, Daniel Rönnlund, Alejandro Fernandez Woodbridge, Noriyuki Sumida, Jerker Widengren, Marta P Imreh, and Rolf Ohlsson. Single cell analysis reveals that H3K9 methylation organizes large, finger-like chromatin hubs radially impinging on the nuclear membrane to protect against large-scale reorganization of chromosome conformations.
Under review in Nature.

* Equal contribution

Related papers not included in this thesis:

I. Anita Göndör, Alejandro Fernandez Woodbridge, **Chengxi Shi**, Xingqi Chen, Moumita Biswas, Erik Aurell, Marta P Imreh, and Rolf Ohlsson. The topology of modular chromosome interactomes impinging on the *H19* imprinting control region is RNA-dependent and physically interconnects loci associated to a wide range of human diseases.
Manuscript.

II. Kuljeet Singh Sandhu, Zhihu Zhao, **Chengxi Shi**, Chandrasekhar Kanduri, and Rolf Ohlsson. Genome-wide chromatin interactions of mouse *H19* imprinting control locus negate a general link to CTCF.
Manuscript.

III. Björn Reinius, **Chengxi Shi**, Liu Hengshuo, Kuljeet Singh Sandhu, Katarzyna J Radomska, Glenn D Rosen, Lu Lu, Klas Kullander, Robert W Williams, and Elna Jazin (2010). Female-biased expression of long non-coding RNAs in domains that escape X-inactivation in mouse.
BMC Genomics, *11*, 614.

CONTENTS

LIST OF ABBREVIATIONS

3C	Chromosome Conformation Capture
4C	Circular Chromosome Conformation Capture
5C	Chromosome Conformation Capture Carbon Copy
5mC	5-methylcytosine
5hmC	5-hydroxymethylcytosine
AS	Angelman Syndrome
BWS	Beckwith-Wiedemann Syndrome
CAP	Chromatin Architectural Protein
CHARM	Comprehensive High-throughput Arrays for Relative Methylation
C-DMR	Cancer cell-specific Differentially Methylated Region
ChIA-PET	Chromatin Interaction Analysis by Paired-End Tag sequencing
ChIP-seq	Chromatin Immunoprecipitation with DNA sequencing
ChrISP	Chromatin *In Situ* Proximity
CT	Chromosome Territory
DMR	Differentially Methylated Region
DNA	Deoxyribonucleic acid
DSB	Double Strand Break
e4C	Enhanced ChIP-4C
EB	Embryoid body
EMI	Electron spectroscopic Imaging
ES cells	Embryonic Stem cells
FISH	Fluorescence In Situ Hybridization
HATs	Histone Acetyltransferases
HDACs	Histone Deacetylases
ICR	Imprinting Control Region
LADs	Lamina-associated Domains
LCR	Locus Control Region
LOCKs	Large Organized Chromatin K9 modifications
LOI	Loss of Imprinting
ncRNA	Non-coding RNA
PAR	Poly(ADP-ribose)
PARG	Poly(ADP-ribose) Glycohydrolase
PARP	Poly(ADP-ribose) Polymerase
PBMC	Peripheral Blood Mononuclear Cell
PcG	Polycomb Group
PGC	Primordial Germ Cell
PWS	Prader-Willi Syndrome
RNA	Ribonucleic acid
RNAP II / III	RNA Polymerase II / III
siRNA	Small interfering RNA
T-DMR	Tissue-specific Differentially Methylated Region
TSS	Transcription Start Site

1 INTRODUCTION

The genomes of more than 200 organisms including human have been sequenced, and the project to sequence 1000 individual human genomes from a number of different ethnic groups was completed in 2012 (The 1000 Genomes Project Consortium, 2012). It has become increasingly clear that inheritable information beyond that embedded in the DNA sequences is pivotally contributing to the complex biology of organisms. This phenomenon, termed 'epigenetics', was initially coined by Conrad Waddington in 1942 as 'the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being' (Waddington, 1942).

Nowadays, in a broad sense, epigenetics studies the causes for the changes of the cellular phenotype without changing the underlying DNA sequences. For instance, in multicellular organisms, the genetic information of each cell is for the most part identical, while the gene expression profiles and cellular functions can be heterogeneous in different cells. This epigenetic regulation of cell fates and tissue heterogeneity has been described as 'epigenetic landscape' by Waddington (Waddington, 1957). Over many years of research efforts aimed at understanding the molecular basis of the epigenetic mechanism, a consensus definition has been arrived, as 'An epigenetic trait is a stable heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence' (Berger et al., 2009).

Recent advances have revealed that the spatio-temporal behavior of genome in the context of nuclear architecture is essential for genomic function regulation. The higher order chromatin conformations and topological organization of genome in three-dimensional space are now considered as the important epigenetic traits. The three-dimensional chromatin and genomic organizations have been shown to be dynamic yet conserved in response to environmental cues. The emerging view is that genome topology is a self-organizing system and there exists self-enforcing feedback between genomic activity and structure.

The work presented in this thesis focused on the study of higher order chromatin conformations and networks regulated epigenetically during development and by environmental cues. The studies also include the invention of a novel method to identify higher order chromatin conformations in single cells with high resolution.

Epigenetics traits and the chromatin structure

There are two main categories of well-studied epigenetic traits: DNA methylation and histone modifications. In the past few years, due to the technique advance, large scale genome-wide analysis of these two epigenetic marks have been achieved with rather high resolution and accuracy. The term 'epigenome' has been used to describe the genome-wide pattern of the epigenetic marks, more specifically 'DNA methylome' or 'histone code'.

DNA methylation

DNA methylation is the only epigenetic mark for which the detailed mechanism of inheritance during mitosis is known in detail (Bird, 2002). In vertebrates, the most common form of DNA methylation is 5-methylcytosine (5mC) (Figure 1), and it typically occurs in the context of CpG dinucleotide. 60%-70% of the CpGs are methylated in mammals. Non-CpG methylation has been observed at low level in embryonic stem (ES) cells and early embryo development in mouse (Haines et al., 2001). Recent studies have shown that other forms of DNA modification, for example, 5-hydroxymethylation (5hmC) is also present in human and mouse ES cells and brains, which is also associated with epigenetic regulation (Ito et al., 2010; Kriaucionis and Heintz, 2009; Pastor et al., 2011; Tahiliani et al., 2009). It has also been linked as a byproduct during active demethylation.

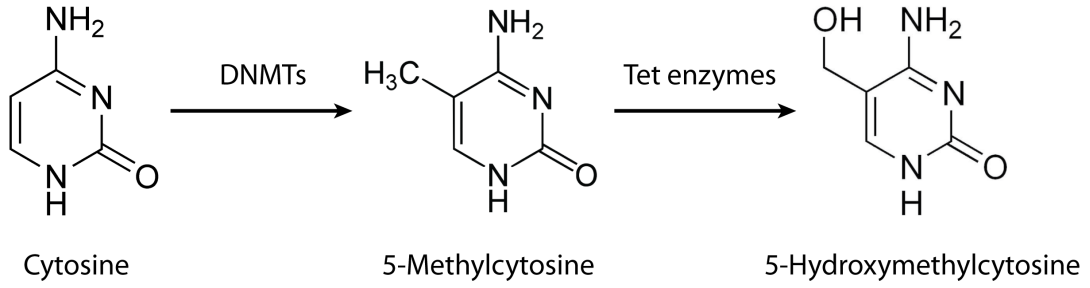

Figure 1. Cytosine is methylated by DNA methyltransferases (DNMTs) into 5-Methylcytosine, 5-Methylcytosine is hydroxylated by Tet family enzymes into 5-Hydroxymethylcytosine.

The CpG dinucleotide cluster together in mammalian genome, and the regions with high (C+G) and CpG content is termed CpG islands (Bird, 2002). The CpG islands only occupy 0.7% of the human genome, but contain 7% of the CpG dinucleotide (Fazzari and Greally, 2004). CpG islands are mostly unmethylated in all somatic cell types at all stages of development, and are associated with around 60% of human gene promoters (Bird, 2002).

DNA methylation pattern is dynamic during development, which can be exemplified by the genome-wide demethylation events initiated during the formation of zygote. Although CpG demethylation is initially an active process for the paternal genome, it is passive for the maternal genome (Reik et al., 2001; Santos et al., 2002). New genome-wide methylation patterns are quickly established at blastocyst stage with ensuing different methylation patterns

specific for different somatic cell types. A small portion of CpG islands is also methylated during development, contributing to the silencing of associated gene promoters. Aberrant DNA methylation pattern may occur in/contribute to certain diseases such as cancer. The genome-wide hypomethylation is a common feature of many types of cancer to cause genomic instability, often associated with local hypermethylation, for instance, on promoters of tumor-suppressor genes.

Over the years, several techniques have been developed to study the cytosine methylation at specific loci, some of which have been combined with high-throughput methods to map the cytosine methylation at a genome-wide scale. (1) Bisulfite-based methods use chemical reaction to convert unmethylated cytosines into uracils to introduce single nucleotide polymorphism; (2) MeDIP-seq uses antibody that specifically binds to 5-methlylcytosine from sonicated DNA; (3) MethylCap-seq employs methyl-binding domain protein to select the methylated DNA; (4) Methylation sensitive digestion uses restriction enzymes to selectively digest only methylated or unmethylated DNA.

These methods have been used to compare genome-wide cytosine methylation patterns (methylomes) in different cell types or in normal and tumor samples. Though they provide a general picture in a larger scale, each of them has its own limitation and bias (Bock et al., 2010).

Feinberg and colleagues have developed an approach termed 'comprehensive high-throughput arrays for relative methylation' (CHARM) by designing original array and new statistical procedures to detect and analyze the fractionated DNA from a methylation sensitive enzyme McrBC (Irizarry et al., 2008). In their following studies they have found that most of the tissue-specific and cancer cell-specific differentially methylated regions (T-DMRs and C-DMRs) are present not in CpG islands but in the sequences up to 2kb distant from CpG islands, which they termed 'CpG island shores' (Doi et al., 2009).

By combining bisulfite methods with high-throughput sequencing, Ecker and colleagues have generated the methylomes in human embryonic stem cells and fetal fibroblasts, which were the first genome-wide, single-base-resolution maps of methylated cytosines in mammalian genome (Lister et al., 2009).

Up to today, complete DNA methylomes for several organisms and different cell types are available (Pelizzola and Ecker, 2011). High-throughput techniques for DNA methylation detection are constantly being improved to become more and more cost-effective. A full understanding of the role of methylome in regulating cellular functions and mechanisms in its establishment and maintenance is now a key question in the field. The emerging methylome data on patient samples will serve for a better understanding of the human diseases, and will be eventually applied for the clinical diagnosis and treatment.

Histone modifications

In eukaryotes, the genome DNA is wrapped around core histones (histone H2A, H2B, H3, H4) to form chromatin. This arrangement enables the very long DNA thread to be packaged in the microscopic nucleus and yet allow developmentally regulated access to key *cis* regulatory elements. As the histone tails can undergo a plethora of post-translational modifications, including acetylation, methylation,

phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, deimination and proline isomerization, the properties of the chromatin can be accordingly modified. The modifications can occur on over 60 residues as determined by mass spectrometry (Kouzarides, 2007) (Figure 2). Though the majority of these modifications are still not well studied, considerable progress have been made on our understanding of acetylation and methylation on histone H3 lysines. Thus, acetylation on histone H3 lysine 9 (H3K9) and methylation on H3 lysine 4 (H3K4) and H3 lysine 36 (H3K36) are normally associated with an open chromatin structure and active state of chromatin. Conversely, methylation on H3 lysine 9 (H3K9), H3 lysine 20 (H3K20) and H3 lysine 27 (H3K27) are generally associated with more compacted chromatin structures and repressed states.

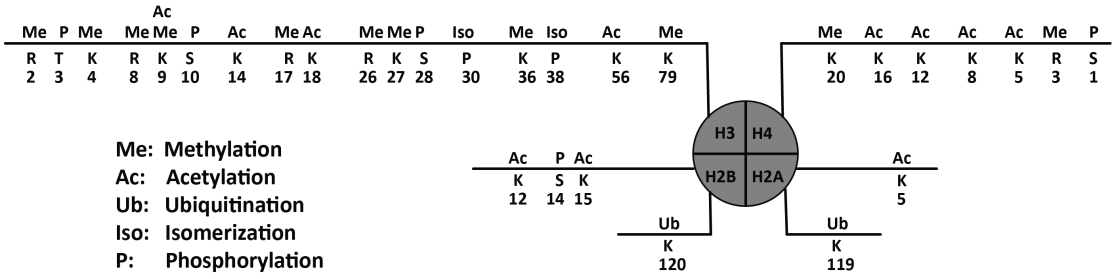

Figure 2. Scheme description of some of the current known modification on core histones.

The modifications on histone tails are not static, but can react to the stimulus from the cell environment within minutes. This dynamic nature reflects an active mechanism to add and remove the modifications. Over the past ten years, many enzymes have been identified to direct the modifications on histone tails, such as for acetylation (Sterner and Berger, 2000), methylation (Zhang and Reinberg, 2001), phosphorylation (Nowak and Corces, 2004), ubiquitination (Shilatifard, 2006), sumoylation (Nathan et al., 2006), ADP-ribosylation (Hassa et al., 2006), deamination (Cuthbert et al., 2004), and proline isomerization (Nelson et al., 2006). Evidences have shown that multiple enzymes are capable to modify the same site on histone, and one enzyme can also modify distinct sites or even some non-histone substrates.

As mentioned above, histone modifications can affect the higher order structure of chromatin, by regulating the contacts between histones or between histone and DNA. For instance, acetylation on lysine can neutralize the basic charge, which will sequentially unfold the chromatin. Modifications on histone tails can also recruit other distinct modification-specific proteins, which may further affect the chromatin structure and accessibility. Proteins recognize and bind to modifications through specific domains. Histone methylation can be recognized by protein carrying chromo-like domains and PHD domains; whereas acetylation can be recognized by protein carrying bromodomains; and phosphorylation can be recognized by a domain within 14-3-3 proteins (Kouzarides, 2007).

The complexity of histone modifications is compounded by their communication with each other, both intra- and inter-molecularly. For example: (1) Distinct modifications on same residue are mutually exclusive. (2) Binding of proteins to a certain modification can be strengthened or disrupted by the nearby

modifications. (3) Proteins recruited to other modified residues can deliver enzymes to modify a certain residue.

The role of histone modifications goes beyond the local context. During the past several years, several reports have described systematic studies of the genome-wide landscape of histone modifications and its regulatory properties to the chromatin (The ENCODE Project Consortium, 2012). Such endeavours have been accomplished by using chromatin immunoprecipitation combined with high-throughput sequencing (ChIP-seq).

Bernstein and colleagues mapped 9 chromatin marks (8 of which are lysine methylation and acetylation) across 9 cell types, to systematically characterize regulatory elements, their cell-type specificities and their functional interactions (Ernst et al., 2011). In 2013, in order to detect the impact of the global chromatin states on cellular phenotypes across different lineages, developmental stages, and environmental conditions instead of *in vitro* cultured cells, the same group mapped 7 histone modifications (H3K4me1, H3K4me3, H3K9me3, H3K27me3, H3K36me3, H3K9ac) for 29 tissues and cell types spanning a wide range of developmental stages, lineages, and derivations (Zhu et al., 2013). The study identified ~ 400,000 cell type specific putative elements; and during development the specification is accompanied by a stark transition in the epigenetic landscape from a uniquely accessible state to increasingly restrictive configurations. Many more studies have employed ChIP-seq approaches to establish datasets of global histone modifications in different cell types in different model organisms, and much more datasets will be obtained. All these datasets enable us to systematically analyze the global chromatin states and their regulation of cellular functions.

The chromatin structure

In eukaryotes, meters of DNA double strands, which serves as the most important information storage needs to be compacted to a large degree and packaged into the limited volume of cell nuclei. This compaction is achieved through several levels of chromatin higher order structure. In interphase nuclei, the chromatin is comparably loose and has different levels of accessibility to proteins, enzymes and other molecules. The metaphase chromatin is highly compacted into separate chromosomes, which in humans have four arm structures visible under light microscope.

The nucleosome string

The nucleosome is the basic unit of chromatin structure, which is composed by an octamer of four core histones (H3, H4, H2A, H2B) wrapped by ~ 147 bp of DNA sequence. The nucleosomes are organized into a 'beads-on-a-string' structure with linker DNA in between, considered as the primary structure of the chromatin. The length of linker DNA can vary between different species, or different cell types in the same organism, or even in the same nuclei (van Holde, 1989). Beside the core histones, histone H1 localizes to the linker region between nucleosomes to presumably participate in the generation of more compact structures associated with repressed states. Histone proteins are among the most conserved proteins in

eukaryotes. In addition, there are several histone variants with very similar core structure to the major ones, but with specialized features. Such minor variants are usually distributed to certain chromatin regions to carry out specialized chromatin functions.

Does the '30 nm fiber' do exist in vivo?

While the 'beads-on-a-string' structure is approximately 10 nm in diameter, the interactions between nucleosomes facilitate the packaging and formation of the larger, secondary structures. The linker histone H1 can bring the entry and exit DNA strands of nucleosome together and play a critical role in promoting the higher order chromatin structure. The organization of the secondary structure can vary due to several reasons: the different length of linker DNA, the core histone variants and the covalent modification on histone tails (Woodcock and Ghosh, 2010). The most direct way to determine the chromatin higher order structure is to visualize the cell nuclei by light or electron microscopy. However, these approaches have resulted in little information obtained due to the very compacted structure of chromatin that is hard to be resolved. The *in vitro* experiments using isolated chromatin have suggested the formation of the chromatin fiber with 30 nm in diameter, known as the '30 nm fiber'. Normal experimental approaches include the digestion of linker DNA by nucleases to generate polynucleosome chains and study their properties under defined conditions in solution. A landmark study has synthesized and crystalized a tetranucleosome chain, and determined the structure by X-ray diffraction (Schalch et al., 2005). Based on these experiments, two models of the nucleosome arrangement in the 30 nm fiber have been proposed: the zigzag two-start helix (Bednar et al., 1998) or the solenoid one-start helix (Robinson et al., 2006). Due to the lack of *in vivo* data, the 30 nm fiber in cell remains controversial. People have argued that it might be merely an artificial structure that appears outside the nucleus system, and the *in vitro* work is 'chasing the mirage' (van Holde and Zlatanova, 1995). Clearly, a more full understanding of this enigma will require the invention of new techniques. A potentially promising approach, under development in several labs, is represented by the thinning of frozen hydrated material with a focused ion beam (Marko et al., 2007). This method does not involve the compression and local heating of cytosectioning, and might provide improved structural information.

The DNA structure beyond the 30 nm fiber

The research work utilizing electron microscopy (EM) coupled with tomography has found fiber-like structure with variety of diameters inside the nuclei (Konig et al., 2007). Thin section observed by EM combined with immuno-gold detection has revealed fibers with diameters from 120 nm to 170 nm (Kireev et al., 2008). The light microscopy techniques are currently limited by the low resolution to resolve the fiber structure. However, live cell imaging can be very useful, if high resolution could be achieved simultaneously, to reveal the real chromatin structure and dynamics *in vivo* without introducing possible structure damage by cell fixation. In recent years, several super-resolution microscopy techniques have been developed which might improve our understanding of the chromatin structure *in vivo*. However, it should be noted that all of these techniques currently

suffer from low resolution in the Z plane to compromise elucidation of the 3D topology of chromatin structures.

Heterochromatin and euchromatin

During the early days of cytology, the area in the interphase nucleus that was stained strongly by basic dyes was defined as 'heterochromatin' (Heitz, 1928). Heterochromatin reflects the more compacted chromatin structure, while 'euchromatin' stained less intensely and is decondensed (McBryant et al., 2006). Heterochromatin and euchromatin are, however, not strict definitions for chromatin higher order structures, but these terms generally reflect the distinct compaction states of chromatin with different accessibility for regulatory proteins and enzymes. Heterochromatin tends to locate at the periphery of nucleus and the surrounding boundary of the nucleolus and replicates late. Euchromatin, on the other hand, is the active portion of the genome, which is generally gene-rich, actively transcribed and replicates early. Heterochromatin can be divided into constitutive heterochromatin and facultative heterochromatin. Constitutive heterochromatin, normally be found at the centromere and telomere area, represents gene-poor regions largely composed of repetitive sequences. Facultative heterochromatin can switch from compacted state to more open and active state in response to environmental cues to contribute to generation of cell type-specific transcriptional repertoires.

Chromatin architectural proteins

As mentioned above, the organization of higher order chromatin structure relies on the interplay within the nucleosome (histones and DNA) and between nucleosomes. However, besides these intrinsic constraints, some proteins that can specifically bind to chromatin are also largely involved in the regulation of chromatin structure, which are known as chromatin architectural proteins (CAPs). CAPs bear totally different features in many ways but only one common characteristic is to insulate architecture of the underlying fiber as the name indicates (McBryant et al., 2006). There are different mechanisms for CAPs to recognize the binding sites on chromatin. For instance, the recognition targets can be specific DNA sequences (Polycomb group (PcG) proteins), DNA methylation (Methyl-CpG binding protein 2 (MeCP2)), and histone modification (HP1) (Woodcock and Ghosh, 2010). CAPs contain one or more chromatin binding domains that enable them to bridge the fibers and compact the chromatin.

Chromatin remodeling complexes, which can be recognized as a special type of CAPs can open up chromatin architecture to increase the accessibility of polymerases and transcription factors for gene activation. The remodeling, which is a dynamic process, is carried out by two major groups of proteins: (1) Histone modifying enzymes can covalently modify the histone tails to regulate the binding affinity between histones and DNA which results in the loosening or tightening of the coil, such as histone acetyltransferases (HATs), histone deacetylases (HDACs), methyltransferases. (2) ATP-dependent chromatin remodeling complexes can use the energy of ATP hydrolysis to move, eject or restructure nucleosomes. These activities can reposition nucleosomes along the DNA strand or replace nucleosome components by histone variants. Currently there are five known families of such remodeler: SWI/SNF, ISWI, NuRD/Mi-2/CHD, INO80, SWR1. They

all share the same ATPase domain, but their functions are involved in different important biological process including DNA replication and repair, apoptosis, chromatin segregation and more.

Epigenetics in 3D

The cell nucleus is a highly organized and complex structure. It was first observed and described by the microscopist Antonie van Leeuwenhoek in 1719. Though it has been thus known for three centuries, the importance of nuclear architecture in the regulation of cellular and genomic functions has started to be unraveled only recently.

The spatio-temporal behavior of the genome in the context of the nuclear architecture plays a critical role in genomic function regulation. This regulation has added up to another layer of three-dimensional epigenetic information that has become under intense focus recently.

The spatial organization of genome

Chromosome Territory

A major step forward for understanding the 3D context of the eukaryotic genome was initiated by Cremer and colleagues who applied fluorescence in situ hybridization (FISH) using specific probes from chromosome sorted by flow cytometry in human lymphocyte nuclei (Rappold et al., 1984). This seminal report provided the first direct visualization of individual chromosome within the interphase nucleus. What their original and subsequent experiments found is that each individual chromosome occupies a certain space inside the nucleus, termed 'Chromosome Territory (CT)' (Figure 3). This territorial organization of genome in the interphase nucleus has been confirmed and accepted as a basic principle of nuclear architecture in both animals and plants (Cremer and Cremer, 2010). The tightly packed CTs intermingle with each other generally on the boundaries in higher eukaryotes (Branco and Pombo, 2006), while in yeast the territories are less well defined and intermix to a larger extent (Zimmer and Fabre, 2011).

Many studies have shown that the chromosome territories are non-randomly arranged in the 3D space with respect to: (1) The radial location in the nucleus; (2) The CT neighborhood; (3) Relative location to other nuclear compartments.

Thus, gene-rich chromosome 19 and gene-poor chromosome 18 in human lymphocyte and other cell types showed that the CT 19 were present mostly in the interior of the nuclei while the CT 18 were found at the periphery (Cremer et al., 2001; Cremer et al., 2003; Croft, 1999). Bickmore and colleagues analyzed the nuclear positions of all chromosomes in human lymphoblasts and primary fibroblasts and confirmed the correlation between the gene density of the chromosome and interior-periphery arrangement (Boyle, 2001). The gene density dependent CT arrangements were also observed in other animal models such as rodents (Mayer et al., 2005), birds (Habermann et al., 2001) and cattle (Koehler et al., 2009). Other factors like transcriptional states, GC content and replication timing have also been shown to be correlated with the radial distribution of chromosomes or sub-chromosomal regions (Federico et al., 2006; Goetze et al., 2007; Grasser et al., 2008; Hepperger et al., 2008; Mayer et al., 2005).

Non-random proximities between subset of CTs have been described in both mouse and human samples. In mouse normal splenocytes, three interphase

chromosomes 12, 14 and 15 are preferentially positioned in close proximity to each other. The same CT neighborhoods were also observed in the corresponding lymphoma cell line, which often displays translocations between regions on these chromosomes (Parada et al., 2002). The correlation between CT proximities and translocation was also observed in human lymphocytes (Brianna Caddle et al., 2007). Moreover, the CT proximities exhibit a tissue-specific pattern, which correlates with the tissue-specific translocation events (Parada et al., 2004). These observations suggest that the non-random chromosomal organization and neighborhoods can determine translocation partners.

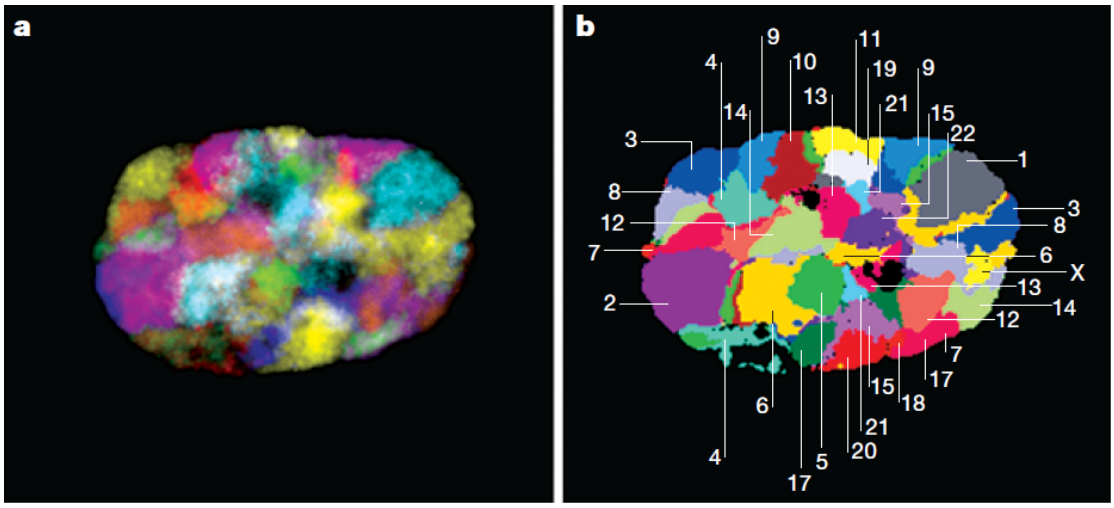

Figure 3. 24-color 3D DNA FISH representation of chromosome territories (CTs) in a human G0 fibroblast nucleus. (a) A RGB image of all labeled CTs (1-22, X and Y) was generated from a deconvoluted mid-plane nuclear section from a 3D stack recorded by wild-field microscopy in eight channels: DAPI, diethylaminocoumarin, Spectrum Green, Cy3, Cy3.5, Cy5, Cy5.5 and Cy7. (b) False color representation of all CTs visible in this mid-section after classification with the software goldFISH. Adapted from (Speicher and Carter, 2005).

3D genomic location

Gene positioning

Studies have shown that in some cases, the radial positioning correlates with the gene activation states with active genes located in the interior and silenced genes at the periphery (Takizawa et al., 2008b). Upon activation, some genes move from a peripheral location to the interior. Such examples include: β-globin genes during mouse erythroid cell differentiation (Ragoczy et al., 2006); *IgH* and *Igκ* during murine B cell differentiation (Kosak et al., 2002); *GATA-3* and *c-maf* during murine T cell differentiation (Hewitt et al., 2004); *Mash1* during mouse neuron differentiation (Williams et al., 2006); *GFAP* activation in murine astrocytes (Takizawa et al., 2008a); *HoxB1* and *HoxB9* in mouse embryos (Chambeyron and Bickmore, 2004). In many other observations, however, the genes did not change radial locations after activation or silencing (Hewitt et al., 2004; Meaburn and Misteli, 2008; Zink et al., 2004). Moreover, genes that relocate to different radial positions during differentiation or pathogenesis do not necessarily change their expression levels (Meaburn and Misteli, 2008; Williams et al., 2006).

It is thus currently not possible to robustly argue that the absolute radial position of gene can serve as the sole regulatory information for gene expressivity *per se*. One reason for this assertion is that the RNA polymerase II (RNAP II) is equally distributed along the radial axis within the nucleus except nucleolus. It seems that the relative position of gene to other landmarks, such as the nucleolus and nuclear envelope, can be the cause of different gene activities. The nuclear envelope has long been speculated as the genomic silent region due to the peripheral distribution of heterochromatin. The artificial tethering of certain chromosomal regions to the periphery led in some, but not all of the cases to down-regulation of the reporter genes (Finlan et al., 2008; Kumaran and Spector, 2008; Reddy et al., 2008). It can be argued that due to the different tethering approaches applied, the association to the periphery does not always bring the genes to the repression machinery present at the nuclear membrane.

Moreover, the gene neighborhood regions may also contribute to the passive change of gene positioning. Responding to developmental or environmental cues, some genes actively reposition for regulated expressivity; while other genes adjacent to them on the linear genome are passively dragged to relocate only as the passenger of the chromatin motion without changing expression level. *Mash1* repositioning and induction during mouse neuron differentiation drags nearby *Pah* and *Igf1* to the interior of nucleus, though the two genes are still silenced (Williams et al., 2006).

Genome - nuclear lamina interactions

Electron microscopy analyses have indicated that some chromatin is in close contact with the nuclear lamina (Paddy et al., 1990). This observation has recently been confirmed: Steensel and colleagues have documented that the genome-lamina interactions occur through more than 1300 sharply defined large domains 0.1-10 Mb in size, termly lamina-associated domains (LADs), in human fibroblasts. These LADs are linked to gene repression and are punctuated by specific regulatory elements at the borders. The existence of LADs may not only serve for the individual gene regulation but also for the organization of overall genome architecture as the anchoring sites of interphase chromosomes to the nuclear lamina (Guelen et al., 2008). Feinberg and colleagues identified large H3K9me2 regions up to few megabases, which largely overlap with LADs in human fibroblasts. They termed these regions 'large organized chromatin K9 modifications (LOCKs)'. LOCKs are mainly gained during ES cell differentiation and exhibit cell type specific pattern; and they are also evolutionarily conserved and linked to gene expression changes in different tissues (Wen et al., 2009).

In our study, we have developed a new high resolution *in situ* method in single cells, to identify and visualize higher order organization of chromatin with H3K9me2 mark and its spatial relationship to the nuclear periphery. This is discussed in Paper IV.

Stochastic, dynamic and semi-conserved genome topology

The topological organization of genome in the three-dimensional space in the interphase nucleus is a fundamental feature of the genome. Based on current evidences, it seems that genome topology information has three main characteristics: stochastic, dynamic and semi-conserved.

The dynamics of genome topology

Live cell imaging has demonstrated that the extent of motion of genes and chromosomes is limited (Cavalli and Misteli, 2013; Chubb et al., 2002), implying that large-scale genome topology is rather stable. However, it has been observed that long-range chromatin movements bring two co-regulated genes together following the stimulus of hormone (Hu et al., 2008; Lin et al., 2009) or virus infection (Apostolou and Thanos, 2008). Live cell imaging analysis has also captured the very brief 'kissing' of two *Xist* regions prior to overt X chromosome inactivation during differentiation of female mouse ES cells (Masui et al., 2011).

Another perspective of the dynamic genome topology comes from the cell type specific organization of the genome. As mentioned above, some genes and chromosome territories undergo repositioning during differentiation to achieve cell type-specific expression repertoires. Also, the genome-nuclear lamina interactions have been identified during subsequent differentiation of mouse ES cells *via* lineage-committed neural precursor cells into terminally differentiated astrocytes. The LADs generally overlap in different cell types with orchestrated reorganization of interaction sites, involving concomitant changes of expression of hundreds of genes during differentiation. This indicates a similar global architecture but different regulation of a set of genes that might be important for cell identity (Peric-Hupkes et al., 2010).

A particularly striking example of large-scale changes in chromatin movement is the global inversion of chromatin structures in mammalian retina (Solovei et al., 2009). Only in nocturnal animals, all heterochromatic regions move to the interior and all euchromatic regions move to the periphery of the nucleus during postmitotic terminal differentiation of rod cell. Surprisingly, due to the higher refractive index of heterochromatin, the reorganized nuclear structure can act as microlenses to help to channel photons to the photoreceptors. This is the first example to show in such a direct way that the cell type specific nuclear architecture significantly contributes to the cellular functions.

The semi-conserved genome topology

An important question on genome topology addresses its heritability. To what extent of the positions and proximity patterns of chromosome territories can be transmitted through mitosis remains controversial. By non-invasive labeling of chromosome subsets and tracking by 4D imaging in rat kidney cells, Ellenberg and colleagues showed that global chromosome positions are heritable during mitosis (Gerlich et al., 2003). In another study, photobleaching experiments have shown that the positions of genes and chromosomes are semi-conserved during mitosis. The overall distribution pattern of a given gene or chromosome in the cell population remains the same, although the individual daughter cells exhibited positional rearrangement after mitosis (Cvackova et al., 2009). However, studies using HeLa cells with GFP-tagged chromatin has shown that CT arrangements are stably maintained from mid G1 to late G2 phase, while major changes of CT neighborhoods occurs though mitosis (Walter et al., 2003). Studies tracking individual loci showed the enhanced chromatin motion and *de novo* establishment of chromatin-nuclear compartments interactions in early G1, suggesting the non-inheritable chromatin positioning (Thomson et al., 2004). At the local level, Ohlsson and colleagues showed that the chromatin loop connecting *H19*

imprinting control region (ICR) and differentially methylated region (DMR) in the 5'-flank of *Igf2* gene mediated by CTCF binding is maintained through mitosis, suggesting the higher order chromatin structure can serve as epigenetic memory during cell cycle (Burke et al., 2005).

The stochastics in genome topology

Though the average patterns of the genome topology are quite constant in a cell population, the individual picture in any given cell can be stochastic. Single cell FISH analysis of genes and CTs showed that not two cells exhibit exactly same genome organization (Cavalli and Misteli, 2013). Indeed, the visualization of the single cell dynamics of genome-nuclear lamina interaction *in situ* revealed that only 30% of LADs are positioned at the periphery in each cell, and LADs are not detectably inherited but stochastically reshuffled upon mitosis (Kind et al., 2013).

To sum up all the current findings, it appears that genome topology is 1) semi-conserved through mitosis; 2) dynamically regulated by environmental cues; 3) exhibits a probabilistic nature and a stochastic pattern in each single cell while the overall distribution pattern in the population remains the same.

Chromosomal interactions

It has long been appreciated that the spatial interactions between regulatory elements contribute to the gene regulation. During the past decade, the development of 'chromosome confirmation capture (3C)' technique (Dekker et al., 2002) and its extensions have offered breakthrough methods to identify the physical contacts between the loci on the same chromosome (*cis*) or from different chromosomes (*trans*).

Intra-chromosomal looping

The interactions between sequences on the same chromosome result in the formation of local chromatin loops. There are four types of loops according to their function in transcription regulation (Cavalli and Misteli, 2013) (Figure 4):

(1) The chromatin loop that brings the promoter and enhancer into contact. The first example was demonstrated by the interaction between β-globin locus control region (LCR) and its promoter (Tolhuis et al., 2002). More such cases have been shown since then. In some cases, the local epigenetic marks such as DNA methylation can mediate to form different local loops that bring different enhancer-promoter combinations together. In the context of mouse and human *H19*/*Igf2* imprinting locus, the imprinting control regions (ICR) on parental alleles are differentially methylated which results in the enhancer interacting with the promoter of either *H19* or *Igf2* on maternal or paternal allele, respectively (Kurukuti et al., 2006).

(2) The second type of chromatin loop is involved in Polycomb-dependent transcription repression. In this scenario, Polycomb bound regions contacts distal gene promoters *via* protein-protein interactions among Polycomb proteins and promoter-associated factors. This type of chromatin loop has been documented in mammals (Tiwari et al., 2008a; Tiwari et al., 2008b).

(3) The 5' of the transcription start site joins the 3' terminal end of the gene. This type of loop has been observed in the case of human rDNA. In this manner, RNA polymerase III will restart another round of transcription in a rapid manner and keep the gene in a high expression level (Nemeth et al., 2008). This type of loop has also been observed on the RNA polymerase II transcribed genes (Tan-Wong et al., 2008).

(4) The fourth type of chromatin loop involves insulator proteins. Insulators often set the boundary between the active and inactive chromatin domains. They do so by connecting two insulator binding sites and forming a loop of the region in between (Hou et al., 2012; Sexton et al., 2012; Van Bortle et al., 2012). The region in the loop is thus separated from the surrounding chromatin environment.

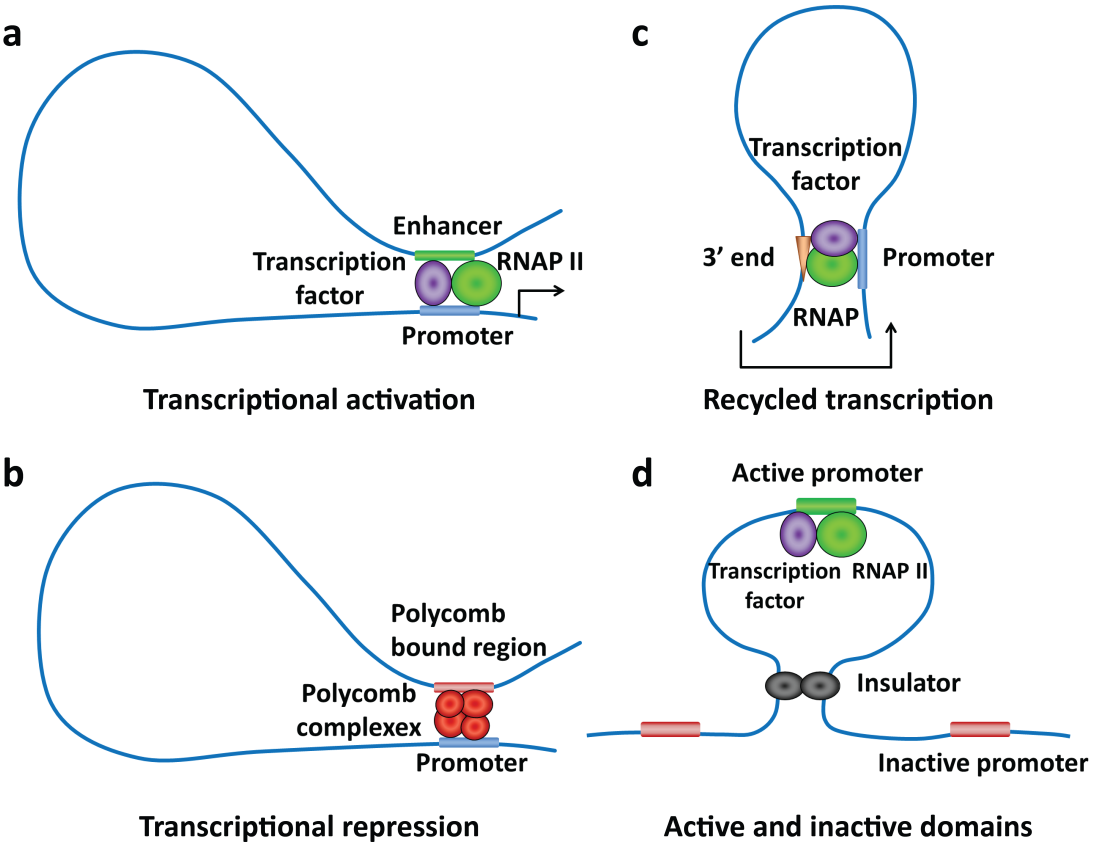

Figure 4. Four types of chromatin loops for transcriptional regulation.

Chromatin loops can bring loci linearly far away into spatial juxtaposition. However, the size of the loop is often restricted by the physical constraints of the chromatin (Göndör and Ohlsson, 2009a). The minimal estimated length for the bending is 0.5 kb for naked DNA and 10 kb for uninterrupted chromatin fibers (Rippe, 2001). Nucleosome free regions such as promoter and enhancer can release the physical constraints on the chromatin fiber thus shorten the possible length for the loops to be constructed.

Inter-chromosomal contacts

While the intra-chromosomal interactions result in the formation of chromatin loops, the contacts between loci from different chromosomes that require larger

chromatin fiber motion have also been widely described by FISH or 3C techniques.

Thus, it has been documented that the T_H2 locus control region (LCR) dynamically interacts either with the promoters of interleukin 4, 5 and 13 (*Il4*, *Il5* and *Il13*) in T_H2 cells, or with the promoter of IFN-γ gene (*Ifng*) in T_H1 cells. The cell type-specific interaction patterns correlate with the cell type-specific gene activation (Spilianakis et al., 2005). One allele of *H19* imprinting control region (ICR) on chromosome 7 physically interacts with one allele of *Wsb1*/*Nf1* on chromosome 11 in a CTCF-dependent manner. Omission of CTCF or deletion of maternal *H19* ICR disassembles this association and alters *Wsb1*/*Nf1* gene expression (Ling et al., 2006). Upon virus infection in HeLa cells, three NF-κB binding regions on different chromosomes interact with the enhancer of *IFN-β* and load transcription factor NF-κB to the enhancer region, thus triggering enhanceosome assembly and *IFN-β* activation (Apostolou and Thanos, 2008). When MCF7 breast cancer cells are stimulated by 17β-estradiol (E_2), ER_α binding sites from different chromosomes interacts with each other in a rapid and specific manner, with a dramatic reorganization of nuclear territories. This rapid motion of chromatin fibers requires the actions of nuclear actin/myosin-I machinery and histone lysine demethylase LSD1 (Hu et al., 2008).

These evidences suggested wide-spread physical interactions between loci from different chromosomes. In many cases, the interactions correlate with gene activation although the relation as cause or consequence is still not very clear.

Chromosomal interaction network

The 3C method is limited to the analysis of two predefined regions (one-to-one), the development of this initial technique has offered the possibility to uncover a wider range of chromatin interactions simultaneously. Circular chromosome conformation capture (Zhao et al., 2006) and chromosome conformation capture-on chip (Simonis et al., 2006) (both termed 4C) can identify genome-wide chromosomal regions that physically interact to a known bait region, which is in a one-to-all manner. Chromosome conformation capture carbon copy (5C) was designed to identify the chromatin interactions in a many-to-many manner (Dostie et al., 2006). The latest derivative of the initial 3C method, termed Hi-C is able to show spatial proximity between any regions to any other regions genome-wide (Lieberman-Aiden et al., 2009). However, the currently low resolution of this technique suggest that it describes more many-to-many rather than all-to-all. The ChIA-PET technique combines the chromatin immunoprecipitation and 3C-based ligation procedure to identify protein binding-based chromatin interactions (Fullwood et al., 2009). All these approaches aim to identify simultaneous multiple chromosomal contacts at the global level and to build up genome-wide chromosomal interaction networks.

Co-regulation of transcription

One of the mostly proposed and observed functional outcomes of the chromosomal interaction networks is to organize and regulate the expressivity of the genome. The transcriptional regulatory feature of the chromosomal interactions can be reflected from several perspectives:

(1) Co-expression of genes

Long before the development of C's techniques, more than two decades ago it has been proposed that transcription might not proceed in the linear way along the genome as in the conventional model, but instead it occurs at certain sites with concentrated RNA polymerases in the nucleus (Jackson and Cook, 1985). Such sites may each contains multiple genes for transcription. The term 'transcription factories' was first coined by Cook and colleagues and used to describe such sites as the basic units for transcription machinery (Iborra et al., 1996; Jackson et al., 1993) (Figure 5). In their pioneered study, Cook and colleagues employed electron microscopy to study nascent transcripts pulse labeled with immuno-gold particles. They found that the nascent transcripts cluster at certain sites with constant size of around 75 nm in diameter. The number and size of the clusters do not change with increase in labeling time. In the same study, RNAP II was also shown to be clustered throughout the nucleus with the size of 56 nm in diameter. The nascent transcript clusters and RNAP II clusters do not totally overlap but with 24nm shift to each other which reflects the structural organization of the transcription factories. Later on in many studies, discrete transcription foci have been observed with in varying numbers in different cell types ranging from several hundreds to several thousands (Jackson et al., 1998; Kimura et al., 1999; Osborne et al., 2004; Pombo and Cook, 1996).

It has been a long-standing mystery about the composition of transcription factories. One study employing electron spectroscopic imaging (EMI) has revealed that transcription factories appear as proteinacious structures with an average diameter of 87 nm, and chromatin and nascent transcripts with RNPs appear as fibrous, phosphorus-rich structures located at the periphery of the factories (Eskiw et al., 2008). In a more recent study, Cook and colleagues isolated the active RNA polymerase complexes in HeLa cells and analyzed the protein contents by mass spectrometry (Melnik et al., 2011). Each complex represents part of the core of the transcription factory, which contains the polymerase subunits, unique transcription factors, nascent transcripts and the DNA templates.

DNA FISH and 3C/4C analyses have revealed many examples that the active genes tend to cluster with other active genes under different scenarios. Using FISH and 3C, Fraser and colleagues have shown that several widely separated genes in a 40 Mb region of distal mouse chromosome 7 dynamically colocalize in shared RNAP II factories in a transcription dependent manner (Osborne et al., 2004). In a later study employing a modified version of 4C (e4C) in mouse erythroid cells, they have shown that the mouse globin genes associated with hundreds of other transcribed genomic loci from different chromosomes in transcription factories. They also showed that the transcription factor Klf1 mediated the co-associated Klf1-regulated genes at a limited number of specialized transcription factories (Schoenfelder et al., 2010). De Laat and colleagues applied 4C to show that the active β-globin locus in mouse fetal liver physically interacts with other transcribed loci on the same chromosome, while the inactive β-globin in fetal brain interacts with transcriptionally silent loci. At the same time, the house-keeping gene *Rad23a* forms interactions predominantly with other active gene clusters in *cis* and in *trans* (Simonis et al., 2006). In a recent study, it has been shown by 4C and ChIA-PET that TNFα responsive genes congregate in discrete transcription factories, which also contains the pathway's downstream

transcription factor NFκB. The author further suggested that genes responding to distinct signal pathways are transcribed in specialized factories, such as 'NFκB factories' for TNFα responsive genes and 'SMAD factories' for TGFβ responsive genes (Papantonis et al., 2012).

All these observations may offer the evidence to confirm that the active genes cluster into the discrete foci for transcription. Each focus may serve as specialized factory with distinct transcription factors for the specificity to recruit certain subset of genes and stability of the complex structure. In this sense, the organization of transcription interactomes can largely contribute to the formation of chromosomal interaction networks.

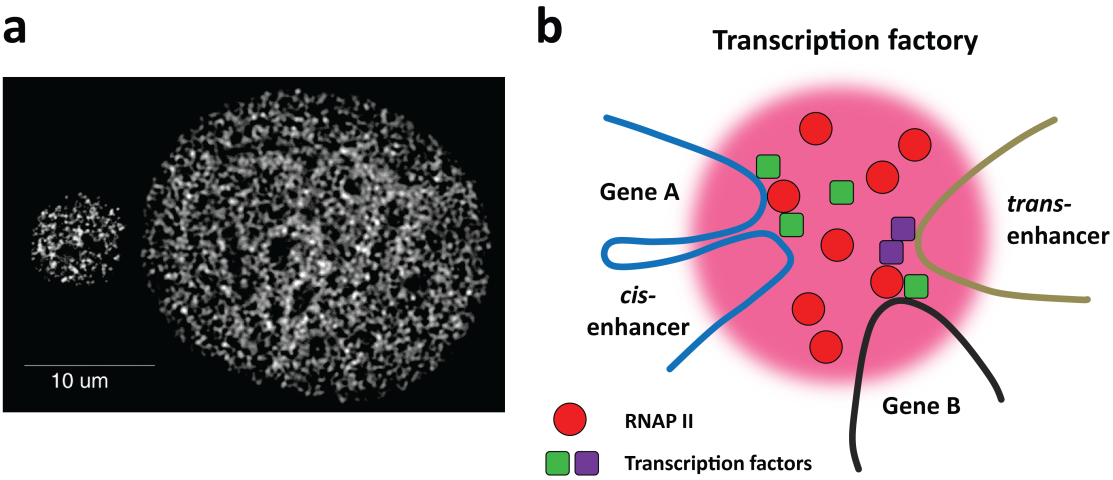

Figure 5. Transcription factory. (a) Maximum intensity projections of Ser5-RNAP II factories in splenic B cell (left) and primary mouse embryo fibroblast (right) nuclei. Adapted from (Chakalova and Fraser, 2010). (b) Model of transcription factory. Genes and distal enhancers from same or different chromosomes converge in the same transcription factory with abundant RNAP II and transcription factors.

(2) Regulatory sequence orchestration

The non-coding part of the genome contains sequence motifs that can control transcription over distance (Splinter and de Laat, 2011). The regulation of transcription over distance can be achieved through the spatial interaction between the gene and regulatory elements in the context of folded genome.

The enhancers and other regulatory elements are likely to play a role in driving the physical clustering of genes (Göndör and Ohlsson, 2009a). Such complexes might include the promoter regions and enhancer elements from the same or different chromosomes. It has been demonstrated that one enhancer can stochastically interact with multiple promoters (Tsytsykova et al., 2007) and multiple enhancers can communicate with one promoter (Deschenes et al., 2007). In one recent study, Dekker and colleagues applied 5C to interrogate comprehensively interactions between transcription starting sites (TSSs) and distal elements in 1% of the human genome representing the ENCODE pilot project regions (Sanyal et al., 2012). They identified more than 1000 interactions

between the promoters and the distal elements resembling enhancer, promoters and CTCF binding sites in each cell line, and the large proportion of the interactions behave in a cell type specific manner. Furthermore, many of the TSSs display more than one long-range interactions - some up to 20; and this multiple interaction pattern is also true for the distal fragments, which interact with up to 10 TSSs. Their study mapped gene-element interactions for a large gene set in a systematical way, and has shown a complex TSS-distal fragments interaction network representing the landscape of chromatin looping. In another study, when human umbilical vein endothelial cells were stimulated by TNFα, the transcription start site (TSS) of TNFα responsive genes *SAMD4A* and *EXT1* converged not only with other TNFα responsive genes but also with large amount of non-genic regions. These non-genic regions were speculated to be enhancers and the data suggested the promoter-enhancer interactions for transcriptional regulation (Papantonis et al., 2012).

Several ChIA-PET studies have provided important insights on the 3D spatial organization and physical contact patterns of the genomic sequences that are enriched for functional or regulatory protein binding. In the ChIA-PET study to elucidate the CTCF binding-chromatin interactome in mouse ES cells, Ruan and colleagues identified *cis*- and *trans*- interactions between gene promoters, regulatory elements and chromatin boundary regions. The data suggests that CTCF can function in genome organization and transcriptional control. Especially, CTCF can regulate gene expression through extensive cross-talk between promoter and regulatory elements (Handoko et al., 2011). In one study focusing on RNAP II-associated long-range chromatin interaction, they found extensive promoter-promoter interactions among proximal and distant genes from 5 human cell lines (Li et al., 2012). This promoter-promoter interaction network does not simply reflect the clustering of expressed genes, since the siRNA knockdown of certain genes can regulate other genes that share the same complex. Taken together, the data suggests that many promoters can cooperatively regulate the activity of other promoters in the multigene complexes. Moreover, promoter-enhancer contacts are also present in the network and the interacting sites are enriched for CTCF, cohesins and chromatin remodeling proteins.

To sum up all the current available evidences, it seems that the intra-genic, extra-genic and inter-genic functional and regulatory elements can spatially communicate and organize into interacting networks for transcription regulation. The physical basis of many of these networks is likely to be the RNAP II-associated complexes with the chromatin regions containing gene promoters and regulatory elements, and proteins to organize chromatin structure and transcription factors. In this manner, the transcription process can be orchestrated coordinately among multiple genes by regulatory elements elsewhere in the genome.

(3) Genes regulated by the same factor

Genes that are co-regulated by the same factors can also associate with each other. Thus, oestrogen administration to human breast adenocarcinoama cells MCF7, induces clusters of most oestrogen receptor α (ER-α) binding sites at gene promoters for coordinated transcriptional regulation. Also it has been speculated that ER-α protein dimers can be recruited to multiple binding sites and interact with other proteins to form chromatin looping structure around target genes; this

conformation may partition genes into sub-compartments in nucleus for transcriptional activation or repression (Fullwood et al., 2009).

In *Drosophila* central brain tissue, PcG target genes interact extensively with each other. Chromosomal interactions between PcG target genes mostly occur on the same chromosome arm suggesting the proximities are also constrained by the overall genome topology but not only PcG binding (Tolhuis et al., 2011). This result is in agreement with the finding that genes of two Hox clusters interact with each other within nuclear PcG bodies in *Drosophila* tissues where they are co-repressed (Bantignies et al., 2011).

Upon T_H2 cell activation, SATB1 is rapidly expressed to organize a transcriptional active chromatin structure of genes located in 200 kb of cytokine locus on mouse chromosome 11. SATB1 binds to multiple SATB1 target sequences (SBSs) to fold chromatin into many small-size loops. The connected SBSs are within promoters and regulatory sequences of *Il4*, *Il5*, *Il13* and *c-Maf*, which are switched on upon T_H2 cell activation (Cai et al., 2006). SATB1 physically connects genes from the same cluster that are co-regulated upon T_H2 cell activation.

Co-association with sub-nuclear compartments

Transcribed genes and other genomic loci can simultaneously localize to the same nuclear body, such as splicing speckles. Co-association of muscle specific genes *myf-4*, *MyoD* and *cMyHC* were found at the SC35 sites in differentiated muscle cells but not in undifferentiated satellite myoblasts or fibroblasts (Moen et al., 2004). In human erythroid cells, it has been shown that five co-transcribed erythroid genes from different chromosomes associate with each other around common nuclear SC35 speckles without sharing the same transcription foci (Brown et al., 2008). The same group has also shown that the association of globin genes at nuclear speckles is not a functional necessity of transcription (Brown et al., 2006). In another human cell line system, the inducible heat shock protein Hsp70 transgenes also associate together with nuclear speckles but not transcription foci (Hu et al., 2009).

It is still debatable that whether the multigene colocalization at certain functional compartments, such as splicing speckles, is merely the probabilistic event due to stochastic sharing of the same machinery, or it is under precise control and can add to another layer of regulation for gene expression.

Coordination of replication

In eukaryotic cells, it has been proposed that the replication origins spatially cluster together to enable synchronous origin firing for a large chromatin domain (Göndör and Ohlsson, 2009b; Pope et al., 2013). Cytogenetic approaches have also shown the discrete replication foci in both yeast and mammal cells. The coordination of replication domains may significantly contribute to the spatial organization of genome and chromosomal interaction networks. This part is discussed in details in the later section on 'replication timing'.

'Junk DNA'

Besides regulatory elements, there is still a large portion of the non-coding part of the genome that contains repetitive sequences, termed 'junk DNA' with unclear functions. However, the vast majority of SNPs identified to be associated with a

range of human diseases map to gene deserts or inter-genic regions, suggesting a functional interpretation of these sequences. In previous studies, repetitive sequences have been shown to be involved in the contacts with other genomic regions. The genomic loci communicating with interferon-β gene in two human cancer cell lines contains Alu elements as a common feature (Apostolou and Thanos, 2008). The mouse *H19* imprinting control region (ICR) interacts with several other imprinted domains surrounded by repeat elements (Zhao et al., 2006).

It has been proposed that repetitive sequences that consist large portion of the genome can set up spatial communication with repetitive sequences located elsewhere, to provide a three-dimensional platform for other functional contacts between chromatin fibers (Göndör and Ohlsson, 2009a). However, due to the technique difficulty to map repetitive sequences, proof in favour of this idea may require the development of new techniques.

Transfer of epigenetic states

Transvection was first coined by E.B. Lewis in 1954 to describe the influence on gene expression by homologous paring. Most of current demonstrated cases are in *Drosophila* where homologous chromosomes are proximal to each other in somatic cells (Duncan, 2002). Allelic transvection describes epigenetic consequences from chromatin fiber interaction such as actions of enhancer or silencer in *trans* between homologous alleles. This phenomenon has also been observed in mouse (Rassoulzadegan et al., 2002) and human (Liu et al., 2008), though with very few cases. By extrapolation, non-allelic chromatin fiber interaction may also able to transfer epigenetic states between the interacting loci (Sandhu et al., 2009). This phenomenon has been shown in our studies and is discussed in Paper I.

Factors regulating chromosomal interaction networks

Many types of factors can mediate physical contacts between chromatin fibers; some of them can even have global function in regulating the chromosomal interaction networks.

As mentioned above, RNA polymerases, transcription factors, activators and repressors can drive the formation of interactions mainly for transcriptional control. Besides these examples, CTCF, cohesin and SATB1 are considered as major architectural proteins in the nucleus. CTCF has been shown to be involved in the chromatin loop formation and gene regulation in the loci of *H19*/*Igf2* (Kurukuti et al., 2006), β-globin (Splinter et al., 2006), antigen receptor (Guo et al., 2011), *HoxA* (Kim et al., 2011), MHC class II (Majumder et al., 2008) and many more examples. There are ~ 40,000 CTCF binding sites in human and murine genome and the protein is highly conserved from *Drosophila* to mammals (Holwerda and de Laat, 2012). CTCF binding sites were found enriched at the borders of LAD domain (Guelen et al., 2008) and the borders of the topological domains identified by Hi-C analysis (Dixon et al., 2012). Such data indicate a global role for CTCF in regulating genome topology organization. In one ChIA-PET study

on the chromatin interactome enriched for CTCF binding, mostly intra-chromosomal and a few inter-chromosomal interactions between CTCF binding sequences were identified (Handoko et al., 2011). The size of the intra-chromosomal loops range from 10-200 kb. Only a few proportion of CTCF binding sites were found to be involved in the chromatin loops and *trans*- contacts formation in this study. Interestingly, the CTCF binding sites are quite conserved between different cell types, so other co-factors may be involved in determining which CTCF sites are engaged in the looping or bridging events. Recently, cohesin has been discovered to bind to CTCF sites in G1 phase of the cell cycle suggesting its additional role besides the function in sister chromatin cohesion (Parelho et al., 2008; Rubio et al., 2008; Wendt et al., 2008). Cohesin is a multiprotein complex forming a ring-like structure to embrace the chromatin fibers, making it a good candidate in influencing chromatin structure organization. Indeed, it has been found to mediate chromatin loops along with CTCF in immunoglobulin locus (Degner et al., 2009), β-globin (Hou et al., 2010), MHC class II (Majumder and Boss, 2011), *H19*/*Igf2* (Nativio et al., 2009), *HoxA* (Kim et al., 2011). Cohesin association to CTCF was found to be cell-type specific in several sites (Chien et al., 2011; Kim et al., 2011), increasing the possibility that CTCF can recruit co-factors like cohesin to mediate cell-type specific chromatin loops. SATB1 is a cell type specific nuclear protein that recruits chromatin-remodeling factors and regulates numerous genes during thymocyte differentiation. It has been shown to form a cage-like structure and selectively tether specialized DNA sequences in thymocyte (Cai et al., 2003).

Non-coding RNAs (ncRNAs) constitutes a large portion of the mammalian transcriptome. Studies have provided evidences to suggest that ncRNAs can contribute to chromatin higher order structure and that ncRNAs form integral components of chromatin fibers. Deep-sequencing of all chromatin-associated RNAs in human fibroblasts has resulted in the identification of more than 200 chromatin-associated ncRNAs (Mondal et al., 2010). Many of the proteins involved in chromatin modifications such as DNA methyltransferases, methyl DNA binding domain proteins (Jeffery and Nakielny, 2004) and HP1 (Muchardt et al., 2002) are capable of binding RNA. Heterochromatin formation in yeast (Chen et al., 2008; Folco et al., 2008) and *Drosophila* (Peng and Karpen, 2007) are also regulated by small RNAs. It has been documented that long ncRNAs can direct and regulate the chromatin activator complexes and chromatin repressor complexes (Rinn et al., 2007; Sanchez-Elsner et al., 2006). At the imprinting loci *Kcnq1* and *Igf2r*, the ncRNAs *Kcnq1ot1* and *Air* have thus been shown to coat the target chromatin regions and interact with histone methyltransferase complexes. The *Kcnq1ot1* ncRNA can recruit the *Kcnq1* locus to a perinucleolar position to mediate the formation of repressive chromatin higher order structures (Nagano et al., 2008; Pandey et al., 2008). In the case of X chromosome inactivation in placental mammals, the to-be inactivated X chromosome transcribes many copies of long ncRNA *Xist* that later coat and compact the entire chromosome leading to its inactivation (Zhao et al., 2008). Considering all these evidences, it is very possible that ncRNAs can mediate the long-range chromatin fiber interaction through direct binding to the chromatin or associated protein complexes. In one of our studies, we have demonstrated the evidence for identifying the molecules to be involved in mediating the chromosomal interaction networks; this topic is discussed in Paper III.

Epigenetics in mammalian development and disease

Genomic imprinting

In the diploid organism, somatic cells contain two copies of all autosomal genes, each of which is inherited from either the mother or the father. While most genes are in general active from both parental alleles in mammals, a small set of genes are expressed mono-allelically in a parent-of-origin manner. This phenomenon, termed genomic imprinting, evolved over 150 million years ago in a common ancestor of marsupials and eutherian mammals (Killian et al., 2000). The parental conflict theory has been proposed to explain the driving force for the evolution of imprinting (Moore and Haig, 1991). The genetic conflict of parental inherited genome is the core of the theory and is based on the promiscuous behavior of early mammals where all siblings could have a different father. The paternal genome evolved a tendency to extract the maternal resources during pregnancy for the offspring fitness, while the maternally expressed imprinting genes limit the fetal growth and nutrition supply in order to divide the resource to more offsprings and maximize maternal reproducibility.

So far more than 100 imprinted genes have been identified in mammals (Barlow, 2011; Bartolomei and Ferguson-Smith, 2011). Although a few these appear as singletons, the majority of imprinted genes are present in clusters containing two or more genes encompassing up to 1 Mb. The expression of imprinted genes is controlled by *cis*-regulatory elements – imprinting control region (ICR), and normally one single ICR controls an entire imprinted cluster. Deletion of ICRs has thus shown to cause the loss of imprinted expression in mouse (Fitzpatrick et al., 2002; Thorvaldsen et al., 1998; Williamson et al., 2006; Wutz et al., 1997). ICRs exhibit parent-of-origin epigenetic modification including DNA methylation and histone modifications. The great majority of ICRs contain parentally differentially methylated regions (DMRs), which can locate within promoter or intergenic regions (Bartolomei and Ferguson-Smith, 2011). The promoter-associated DMRs are methylated on the maternal, but not on the paternal allele for all the currently identified cases, while the DMRs in intergenic regions in four cases (*H19*/*Igf2*, *Dlk1*/*Dio3*, *Rasgrf1*, *Zdbf2*) shows a preference for paternal-specific methylation.

Long-range chromatin interactions have been demonstrated to be able to serve for the imprinted expression. In the case of *H19*/*Igf2* cluster, the *Igf2* and *H19* genes are 90 kb apart on the distal end of mouse chromosome 7 and human chromosome 11. The *H19* ICR is located 2-4 kb in the 5'-flank of the *H19* gene. Loop formation mediated by CTCF/cohesin binding to maternal unmethylated ICR restricts *Igf2* promoter in the inactive domain away from enhancer. On the paternal allele, the distal enhancer interacts with the *Igf2* promoter to trigger its transcription while the paternal *H19* gene is repressed (Kurukuti et al., 2006; Murrell et al., 2004). The chromatin loops remain through mitosis to potentially serve as an epigenetic memory (Burke et al., 2005).

A key question is concerned with the establishment and maintenance of parental different epigenetic marks on ICR. Analysis of primordial germ cells

(PGCs) shows that the methylation patterns of imprinted genes are similar to somatic cells during embryonic migration (Hajkova et al., 2002). Following migration of PGCs to genital ridges, the DNA becomes demethylated (Hajkova et al., 2008). The remethylation on ICRs will occur at different time points in male or female germ cells. In male germline, some ICRs are remethylated in prospermatogonia or spermatocytes (Davis et al., 1999; Li et al., 2004). Remethylation on other ICRs in femal germline occurs after birth during the oocyte growth phase prior to ovulation (Lucifero et al., 2002). All ICRs tested except *Rasgrf1* employed *de novo* DNA methyltransferase DNMT3A and its stimulatory protein DNMT3L for remethylation in germ cells (Bourc'his and Bestor, 2004; Bourc'his et al., 2001; Hata et al., 2002; Kaneda et al., 2004). Following establishment, the imprint pattern must escape the extensive post-zygotic epigenetic reprogramming. DNMT1 is present at very low level in preimplantation embyos and is the most possible candidate for this task (Cirio et al., 2008; Hirasawa et al., 2008; Howell et al., 2001). Other factors have also been identified to be involved in this process such as ZFP57 and PGC7/STELLA (Li et al., 2008; Nakamura et al., 2007).

Perturbation of genomic imprinting is associated with cancer and other complex diseases. Small deletions were found in the imprinted regions on human chromosome 15q11-q13 in the patients with Prader-Willi syndrome (PWS) and Angelman syndrome (AS) (Horsthemke and Wagstaff, 2008). Beckwith-Wiedemann syndrome (BWS) was found to be caused by the epimutation of DMR on *IGF2*/*H19* or *KCNQ1*/*CDKN1C* both on chromosome 11p15 (Weksberg et al., 2005). Loss of imprinting of *IGF2* has been shown to increase IGF2 efficiency of signaling (Kaneda et al., 2007) and be associated with several types of cancer (Sakatani et al., 2005).

Replication timing

Replication timing as epigenetic mark

DNA replication in every cell cycle is the basis for biological inheritance. In eukaryotic cells, the genome contains large chromosomes with tens of thousands of replication origins. The temporal order of replication of large segments of genome during S phase is under precise control. (Pope et al., 2013). Recent studies have indicated that replication timing is inheritable information and serves as an epigenetic mark (Göndör and Ohlsson, 2009b; Hiratani and Gilbert, 2009).

Large domains with synchronous replication timing have been observed correlating with the cytogenetic data about ‘replication foci’ (discussed in the next section). The replication domains vary in size from few hundreds kb to several Mb. When either mouse or human ES cells are differentiated into any of the several lineages, global changes of replication timing affecting around 20% of the genome were detected between any cell type pair comparison. Genetically polymorphic ES cells and respective induced pluripotent stem cells (iPS cells) showed almost identical replication timing pattern. When the replication timing profiles of 22 cell lines (can be divided into seven developmental states during mouse embryogenesis) were analyzed for hierarchical clustering, the cell lines could be

clustered into seven groups, which correlates well with the developmental states (Hiratani et al., 2010). Close to 50% of the genome exhibits reorganization in replication timing at some stage of development. Interestingly, the replication timing patterns of the same cell type from mouse and human origins show strong correlation, which indicates an evolutionary conservation (Ryba et al., 2010). These studies offered the evidences that the replication timing is a conserved epigenetic signature of cellular differentiation state.

It has been noticed for decades that there exists a close relationship between replication timing and transcriptional activity. Early replicating DNA coincided with GC-rich, R-bands with high transcriptional activity, while late replicating DNA coincided with AT-rich, G-bands with low transcriptional activity (Stambrook and Flickinger, 1970). Recent comparison of transcription microarray and replication timing microarray data demonstrated a strong positive correlation (Gilbert et al., 2010). Approximately 75% of the active genes replicate in the first half of S phase and most non-genic regions replicate in late S phase (Hiratani et al., 2008). Upon differentiation, many up- and down- regulated genes exhibit early-to-late or late-to-early replication timing changes respectively. However, this correlation is not absolute for all the genes with some exceptions. This may due to the incomplete transcription arrays lacking noncoding transcribed regions. Another possibility is that for many genes, the local sequence information such as strong nucleosome-positioning sequences or CpG islands may override the influences of replication timing (Gilbert et al., 2010).

Besides transcriptional activity, the relationship between chromatin states and replication timing has also been examined in mouse ES and neuronal precursor cells by the same research group (Hiratani et al., 2008; Ryba et al., 2010). A positive correlation between early replication timing and active chromatin marks (H3K4 methylation, H3K9 acetylation, H3K36me3) was observed, while the correlation is not convincing for late replication timing and repressive chromatin marks. Again, more precise ChIP-seq data for a larger range of chromatin parameters may be needed before drawing more robust conclusions.

Spatial organization of replication

In eukaryotic cells, synchronous firing of spatially clustered replication origins enables the replication of large domains within a certain portion of S phase (Cook, 1999; Göndör and Ohlsson, 2009b; Pope et al., 2013). Cytogenetic approaches by pulse-labeling with nucleotide analogs visualize the clustered replication sites known as ‘replication foci’ or ‘replication factories’ (Berezney et al., 2000; Jackson and Pombo, 1998; Wu et al., 2006). Labeling with fluorescent nucleotide or replication fork proteins can provide the live cell imaging of the replication foci (Leonhardt et al., 2000; Manders et al., 1999; Sadoni et al., 2004). In budding and fusion yeast, there are generally 15-30 foci and they are in a mobile manner to fuse with other foci or split into new foci (Kitamura et al., 2006; Meister et al., 2007). The replication foci in mammalian cells are around 5,000 – 10,000 and not as mobile as in yeast (Maya-Mendoza et al., 2010). Each focus may contain 6-20 replicons and can replicate around 1 Mb in 45-60 minutes. Labeled foci remain in the next several cell cycles, indicating that the replication foci are stable structural units (Jackson and Pombo, 1998).

Cytogenetic experiments also showed the different patterns of replication foci

distribution during S phase. Early replication through the first half of S phase occurs in the interior of the nucleus; the middle S phase replication takes place at a more peripheral location; replication in late S phase replicates large blocks of heterochromatin at the nuclear membrane. Such spatio-temporally regulated replication foci may be the cytogenetic equivalents of replication domains identified by microarray studies that also coordinate synchronously fired replication origins.

This possibility highlights the spatial compartmentalization of chromatin regions to form replication foci and replicate at different time during S phase (Gilbert, 2001; Göndör and Ohlsson, 2009b; Ryba et al., 2010) (Figure 6). This spatial compartmentalization may be facilitated by the physical contacts between chromatin fibers or in turn promote such interactions. Gilbert and colleagues compared their genome-wide replication timing dataset with the genome-wide chromatin interaction dataset from Hi-C experiment in mouse lymphoblast and found a striking positive correspondence (Ryba et al., 2010). The replication domains strongly correlate with the interaction compartments in a cell type-specific manner. This result indicates that sequences localize near each other replicate at similar time. Following completion of the S phase, the replication timing is reestablished in each cell cycle coincident with the anchorage of chromosomal regions and self-assemble of sub-nuclear domains in early G1 phase. The authors proposed a model that describes spatially separated 'fractal globules', which are the basic organization of genome identified by Hi-C, as equivalent to temporally separated replication domains. Later, Arneodo and colleagues employed 4C in cycling and resting peripheral blood mononuclear cells (PBMC) and identified megabase-size self-interaction domains. The boundaries of these domains coincide with early-initiation zone in every cell types. The presence of replication timing related domains even in resting cells indicates the replication-related chromatin organization is maintained beyond mitosis (Moindrot et al., 2012). The relationship between replication timing and 3D genome topology suggests the temporal regulation of replication may achieved through spatial domain organization of the genome. Moreover, replication timing profiles may provide a readout of chromosomal interactions and a mean to identify 3D organization changes during differentiation.

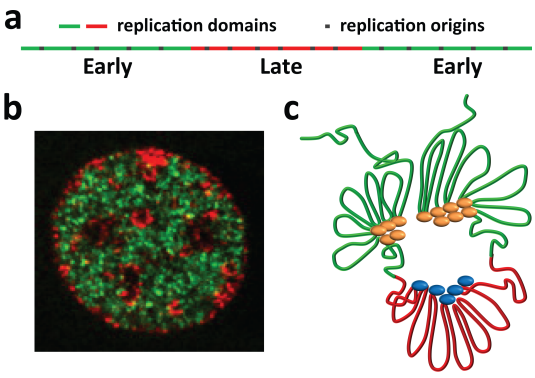

Figure 6. Spatial organization of replication. (a) Large replication domains with synchronous replication timing containing multiple replication origins. (b) Early and late replication foci in mouse embryonic fibroblast nucleus. Mouse embryonic fibroblasts were labeled for 10 min with CldU, chased for 5 h, labeled for 10 min with IdU, and stained with fluorescent antibodies specific to CldU (green) and IdU (red). Adapted from (Wu et al., 2006). (c) Coordination of replication timing might involve origin clustering at different sub-nuclear locations, representing the replication foci in (b).

Asynchronous replication timing of imprinted loci

Genomically imprinted regions undergo asynchronous replication timing during S phase. In most of the tested cases, the paternal allele replicates early in S phase while maternal allele replicates late (Kitsberg et al., 1993). The pattern of replication timing for imprinted loci is erased before meiosis in germ line, and reset to the parental-specific pattern during gametogenesis (Simon et al., 1999). For instance, the late-replicating *Igf2* allele (previous maternal allele) shifts to early replicating during spermatogenesis; the early-replicating *Igf2r* allele (previous paternal allele) shifts to late replicating during oogenesis.

One obvious question is concerned with the mechanisms regulating the asynchronous replication timing patterns. Jaenisch and colleagues showed that the asynchronous replication timing of imprinted loci is unaffected in Dnmt1- and Dnmt3L-deficient ES cells without genomic imprinting. In the case of *H19*/*Igf2* locus, it was argued that the *H19* ICR regulates the allelic-specific sub-nuclear location and replication timing of the locus (Gribnau et al., 2003). Ohlsson and colleagues have shown that by using mouse strain harboring mutant maternal CTCF binding sites on *H19* ICR, the maternal *H19*/*Igf2* locus undergoes a late-to-early switch of replication timing in embryonic liver and brain cells (Bergström et al., 2007). CTCF exerts epigenetic controls over this imprinted domain through insulation activity, transcriptional activity and local higher order chromatin structure, some mechanisms of which might be involved in the replication timing regulation. CTCF binding on *H19* ICR also exhibits *trans*-regulation over other imprinted loci in the genome through chromatin fiber interaction, which is discussed in Paper I.

Chromatin interaction and chromosomal translocation

Chromosomal translocations and corresponding gene fusion products have important roles in tumorigenesis and they account for ~ 20% of human cancer morbidity (Mitelman et al., 2007). Most translocations occur by mis-joining the ends of double strand breaks (DSBs). Two mechanistic factors can contribute to the frequencies of translocation events: the frequency of DSB and the 3D spatial proximity between two chromosomal regions.

In the past few years, several studies have demonstrated the contribution of higher order genome organization to the common recurrent translocations in cancer cells. *MYC*, *BCL* and immunoglobin loci, which are recurrently translocated in different B-cell lymphomas, are preferentially proximate to each other in space in normal B cells (Roix et al., 2003). In splenic B cells deficient of non-homologous end-joining, frequent translocations occur between *Igl*-*Igh* and *Igl*-*Myc* due to the accumulation of DSBs. Correspondingly, *Igl* gene frequently co-localize with *Igh* and *Myc* in these cells (Wang et al., 2009). It has also been shown that tissue-specific translocations correlate with the tissue-specific spatial proximity between chromosomes (Parada et al., 2004), which further illustrates the influence of genome organization in chromosomal translocations.

Very recently, genome-wide correlation between translocations and chromosomal interactions has been examined. Dekker and colleagues compared

the Hi-C genome spatial organization map and genome-wide translocation sequencing dataset in mouse pro-B cells. To directly access the contribution of spatial proximity, they saturated the genomic DSBs by ionizing radiation. Translocations are enriched in *cis* along each chromosome and also to some degree in *trans* directly correlating to the pre-existing spatial proximity (Zhang et al., 2012). Casellas and colleagues have also shown in the absence of recurrent DNA damage in B cells, translocations between *Igh* or *Myc* and all other genomic regions are directly related to their contact frequency by comparing the 4C interaction map and genome-wide translocation map. Again, most rearrangements occur intra-chromosomally although inter-chromosomal rearrangements were also observed. (Hakim et al., 2012).

Interestingly, the rearrangements in *trans* in Casellas' study are biased towards transcriptionally active genes. *Igh* and *Myc* are the most frequently translocation partners in plastmacytoma and Burkitt lymphoma. This likely reflects the observation that *Igh* and *Myc* share the same transcription factories (Osborne et al., 2007). Similarly, *TMPRSS2* and *ERG* or *ETV1*, which all contain binding sites for the androgen receptor, could be induced to translocate in prostate cancer cells treated with androgens (Lin et al., 2009; Mathas and Misteli, 2009).

The accumulated evidences suggest that genome organization and spatial proximities constitute the major driving forces for recurrent chromosomal translocations. Cellular functions like transcription that brings distal genes into the same machinery might significant contribute to this event. Recurrent translocations can thus be the reflection of transcription factor binding patterns. Cell type-specific transcription factor binding patterns and 3D genome topology may offer a plausible explanation for the cancer type-specific recurrent translocations.

2 AIMS OF THE THESIS

The overall purpose of the studies presented in this thesis was to study the nature of higher order chromatin conformations and networks and their developmental regulation in mouse and human model system. Further, the studies also aimed to invent novel techniques that enable the visualization of higher order chromatin proximities in single cells at a high resolution. More specifically, this thesis aimed to:

- To identify chromosomal interaction networks from the window of *H19* imprinting control region in both mouse and human. (Papers I, II, III)
- To examine whether or not chromosomal interaction networks rewire during development. (Paper I)
- To study the epigenetic perspective of chromosomal interactomes and higher order chromatin conformation. (Papers I, IV)
- To study the influence of chromosomal interactomes on transcription and transfer of epigenetic marks. (Paper I)
- To examine the re-organization of chromosomal interactomes in response to environmental cues. (Paper III)
- To identify the possible molecular entities that mediate chromosomal interaction networks. (Paper III)
- To develop a high-resolution, single cell method to visualize chromatin fiber proximity and higher order chromosome conformation. (Paper IV)

3 RESULTS AND DISCUSSION

H19 imprinting control region (ICR) as the bait for 4C analysis

The *H19* and *Igf2* imprinted genes are located at the distal end of chromosome 7 in mouse and distal portion of the short arm of chromosome 11 in human. The imprinting control region (ICR) is positioned in the 5'-flank of the *H19* gene to control their parent-of-origin expression patterns. The mechanism underlying this feature is differential DNA methylation: The unmethylated maternal ICR allele binds CTCF and form chromatin loops that isolate the *Igf2* promoter from the downstream enhancer. Conversely, the paternal ICR allele is methylated rendering the CTCF binding sites inaccessible for CTCF. This is in turn allows the physical contact between enhancer and *Igf2* promoter to generate an active paternal *Igf2* allele (Figure 7).

The loss of imprinting (LOI) at the *H19* ICR leads to not only biallelic *Igf2* expression, but also, to sensitization to IGF-II signaling (Kaneda et al., 2007). Deletion of maternal ICR in mice predisposes to colon cancer (Sakatani et al., 2005) and facilitates parthenogenesis (Kono et al., 2004). Accordingly, LOI is common in human cancers with a maternal to paternal switch in epigenotype. It is reasonable to assume, therefore, that the *H19* ICR has more global influences over the genome separate from its function in the local context. We thus chose *H19* ICR as the bait in many of our 4C analyses to uncover its genome-wide interacting partners and the possible functionalities conveyed by such physical communications.

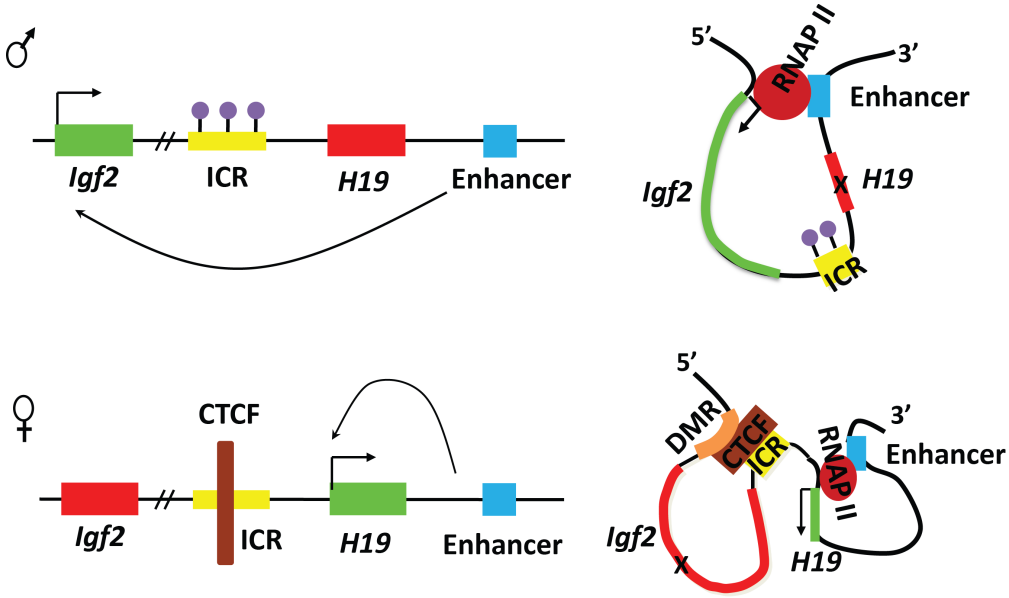

Figure 7. Higher order chromatin structure of *H19*/*Igf2* imprinted domain. The enhancer on paternal allele interact with the *Igf2* promoter; CTCF mediates chromatin looping on the maternal allele to isolate *Igf2* into inactive domain.

Paper I

H19/*Igf2* domain is located at the distal end of chromosome 7 in mouse. The parental alleles of *H19* imprinting control region (ICR) are differently methylated, which controls the expression states of these two genes. In this study, we performed 'Circular Chromosome Conformation Capture (4C)' assay to identify genome-wide interacting partners of *H19* ICR. To achieve a comprehensive screening, 4C DNA samples of mouse neonatal liver, neonatal brain, embryonic stem (ES) cells, and derived embryoid body (EB) cells were pooled and hybridized to whole genome tile path microarray with 100 bp resolution. All captured sequences considered as potential interactors were included in the production of dedicated microarrays to screen for global interaction patterns during ES-EB cell differentiation.

Rewiring of the chromosomal interactome during mouse ES cell differentiation

In the 4C screen, the ES cells displayed a significant loss of intra-chromosomal interactions during differentiation to EB cells. Conversely, the pattern of inter-chromosomal interactions underwent an almost complete reprogramming. These results suggest that the genome-wide chromosomal interactome impinging on the *H19* ICR bait is sensitive to changes in epigenetic states during mouse ES cell differentiation. This possibility may reflect the *H19* ICR is mainly located inside or on the edge of chromosome 7 territory in ES cells, while it more frequently loops out from its own territory in EB cells. The different chromatin environments of *H19* ICR may thus influence the chromosomal regions that it physically contacts with. I conclude that a particular region may regulate *cis*- or *trans*- interactions *via* its own dynamic movement.

Imprinted interaction network and its dynamic pattern

The 4C library revealed a significant overrepresentation of known imprinted genes from 13 chromosomes. As most of these imprinted interactors could also be confirmed in EB, ES and neonatal liver cells, a widespread interactions between *H19* ICR and other imprinted domains in several cell types independent of differentiation status and expression levels of the involved loci could be inferred. We applied 3D DNA FISH to quantitatively analyze the all-to-all interactions among imprinted loci from 7 chromosomes. The result revealed a crosswise imprinted interacting network among all these loci in ES cells. The interactions occur in a one-to-one pairwise manner that we termed 'date' interaction, instead of 'party' interaction forming by more than two imprinted domains simultaneously collide together. This indicates a very dynamic nature of the interactions, which may only last for a short period in every cell cycle. Also as expected, the *cis*- interactions occur inside chromosome 7 territory, while *trans*-interactions are mostly located at the edge of territory. This result further reinforced the notion that the bait *H19* ICR can regulate *cis*- and *trans*- interactions by its own dynamic movement.

Re-organization of imprinted interaction network during epigenetic reprogramming

To examine the influence of epigenetic reprogramming on imprinted interaction network, we turned our focus to male germline development. The paternal specific methylation pattern of *H19* ICR on maternal allele is acquired during spermatogonia-to-spermatocyte transition. Interestingly, the close physical proximities between the imprinted domains could be documented in spermatogonia, but not in spermatocytes or round spermatids, representing two cell types with completed reprogramming of the bait. This result raised the possibility that the imprinted interactome is impinging on the unmethylated maternal *H19* ICR when complexed with CTCF. To address this possibility, we examined spermatogonia harboring a maternally inherited, mutant *H19* ICR allele unable to interact with CTCF. Indeed, all the interactions between the bait and imprinted interactors were absent. Strikingly, even the interactions between the imprinted domains other than the bait were also lost suggesting the CTCF binding on wild type maternal *H19* ICR is essential for the imprinted interaction network during male germline development. We speculated that it is the maternal but not the paternal inherited genome that governs the physical network between imprinted chromatin fibers.

Non-allelic transvection of epigenetic states

To identify the functional outcome of the imprinted interaction network, we exploited previous observations that imprinted domains display asynchronous replication timing pattern, i.e. the parental alleles replicate at different times during the S phase. Moreover, we have earlier documented that the CTCF binding sites of maternal *H19*/*Igf2* allele regulate replication timing of the entire imprinted domain. By extrapolation, the *H19* ICR might regulate replication timing of the regions it interacts with. That turned indeed out to be the case. Thus, the asynchronous replication timing of 7 interacted imprinted domains were switched to synchronous in spermatogonia when the mutant *H19* ICR allele was inherited maternally. The analysis in mutant neonatal liver cells further showed that it is the late replicated allele switches to early replication for all 7 interacted imprinted domains (early/late -> early/early pattern). The mutant ES cells that were used for generating the mutant mouse strain did not appear to be affected by the lack of CTCF binding sites on its maternal *H19* ICR allele indicating that these effects required germline transmission. We concluded that the CTCF binding sites within *H19* ICR can regulate the replication timing pattern of imprinted domains that *H19* ICR physically interacts with during male germline development; and the physical network of chromatin fibers can be utilized as the mean to transfer such an epigenetic mark, i.e. non-allelic transvection.

Summary and perspective

By employing the Circular Chromosome Conformation Capture (4C) technique followed by microarray analysis of the amplified sequences, we identified a developmentally regulated, genome-wide interactome impinging on the *H19* ICR in mouse. The dynamic spatial locations of *H19* ICR taken together with local chromatin conformations likely contributed to its patterns of interactions with other regions in *cis* and in *trans*. We also identified an imprintome, which

exhibited a pairwise pattern of interactions. The imprintome was disassembled during epigenetic reprogramming during male germline development suggesting the existence of a network involving primarily maternal alleles. Importantly, we showed that CTCF binding sites on the maternal *H19* ICR allele were able to regulate in *trans* replication timing patterns to regions it interacted with. We further hypothesized that this imprintome might be involved in the evolution of imprinted states genome-wide. Indeed, the *H19* ICR is the oldest known imprinting control region and the *H19*/*Igf2* cluster is imprinted already in marsupials while most other imprinted genes are singletons (Figure 9). Taken together with our data, it is reasonable to assume that the *H19* ICR has on an evolutionary scale transferred asynchronous replication timing patterns both in *cis* and in *trans* to establish imprinted states.

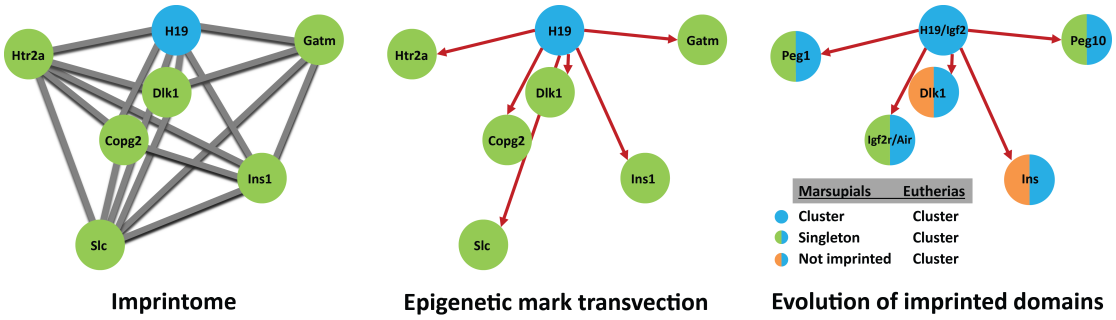

Figure 8. The physical interaction network – imprintome – among imprinted domains in mouse is dependent on CTCF binding sites on the maternal *H19* ICR allele. The imprintome functions as a vehicle to transfer epigenetic states from *H19*/*Igf2* domain. The *H19* ICR might establish imprinted states to other domains through physical interactions during evolution.

Paper II

Chromosomal interactome beyond H19 ICR

In this study, we applied 4C-seq using *H19* ICR as the bait in human ES and EB cells. When we compared the interactome impinging on *H19* ICR to the one in mouse from Paper I, less than 10% could be identified as conserved interactors (not shown). This observation likely reflects evolutionary changes in chromosomal neighborhoods to modify the character of the interactome. The fact that more than two chromosomal sequences occasionally converged in the same complex with physical proximities offered the possibility to capture multiple interactions simultaneously. The ligated circular DNA often contains sequences from more than two genomic regions. We used 454 sequencing to be able to obtain long sequence reads providing a higher chance to map multiple interactors. Thus, the sequencing data provided not only interacting partners of *H19* ICR, but also the interaction between the captured regions themselves in the presence of the ICR.

Modular chromosomal interaction network

From this three-way sequencing data, we established an interaction network with crosswise connections among each node (interactor). Each node has different levels of connectivity depending on the number of other interactors it encounters with (Figure 9). Based on the connectivity of each node, we can construct the topology of the interaction network to observe a highly structured modular network with several central hub-nodes connected to each other and to other connectors and outliers (unpublished). The central hubs reflect a hierarchal structure and the analysis shows the network is scale-free. We also performed 3D DNA FISH to confirm the modular topology of the network in ES and EB cells.

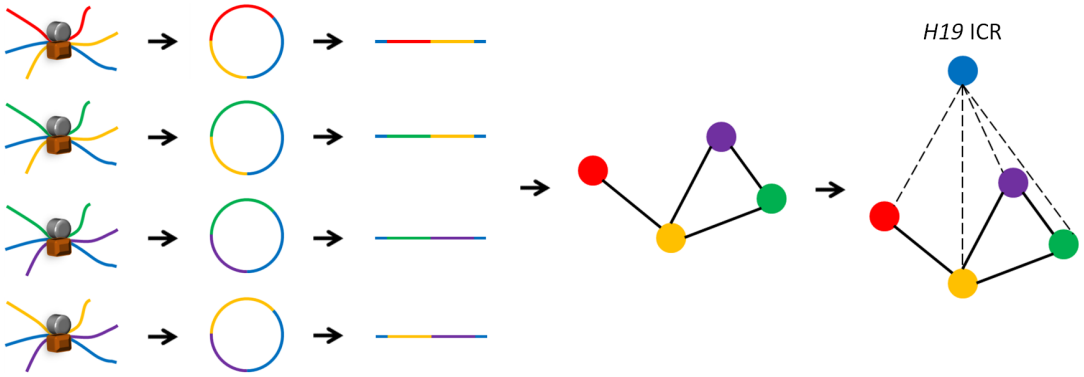

Figure 9. Illustration of 4C-seq technique to capture multiple physical proximities among DNA sequences simultaneously. The three-way sequencing data enables the construction of interaction network with crosswise connections among each interactor.

Summary and perspective

Using 4C-seq in human ES and EB cells, we have identified chromosomal interactome to *H19* ICR, which is not conserved when comparing to *H19* ICR interactome in mouse. The ability to capture interactions between ICR-interacting partners enables us to construct a chromosomal interaction network with crosswise connections. We have further shown the network has a modular topology and is scale-free. The visualization and presentation of topology based on the node connectivity may offer a new perspective for understanding the chromosomal interaction network. The chromosomal regions sharing the same module may also present in the same functional group at any level due to their extensive connectivity among each other.

Paper III

In this study, our main aims were to identify the possible molecules that can mediate the physical interactions between chromatin fibers and to uncover the re-organization of chromosomal interaction network responding to external signaling.

We employed the 4C-seq method to identify genome-wide chromosomal regions physically interacting with the *H19* ICR on chromosome 11 in human ES (HS181 line) cells and derived embryoid bodies (EBs). As a result, we captured in total 507 regions together in ES and EB cells at 1 Kb resolution. The majority (>90%) of these interactors represents inter-chromosomal contacts.

Poly(ADP-ribose) as the mediator of chromosomal interaction network

Our lab has previously documented that poly(ADP-ribose) (PAR) associated with CTCF bound to maternal *H19* ICR is essential for long-range chromatin insulation. Based on this information, we speculated if PAR molecules at *H19* ICR or associated with other chromatin regions might contribute to the establishment of chromatin fiber communications. To answer this question, we treated the already crosslinked chromatin from both ES and EB cells with PAR glycohydrolase (PARG) to digest PAR chains. Then we performed 4C-seq on the treated samples and compared with the data from control samples. A striking disassembly of the majority of the interactors to *H19* ICR was observed upon the removal of PAR. We were able to experimentally exclude the possibilities of unintentional RNA degradation as an unwanted cause of these results. As the presence of the cohesin ring at the *H19* ICR was unaffected by the inhibition of PARP activity, we concluded that PAR serves as a glue molecule to bridge chromatin fiber communications in human ES and EB cells.

TGFβ disconnects Poly(ADP-ribose) dependent chromosomal interaction network

In previous studies from our lab, it was shown that CTCF bound to maternal *H19* ICR interacts with SMAD2/3/4, which are the downstream transducers of TGFβ signaling. Interestingly, our 4C experiments revealed that the *H19* ICR physically interacts with *CHD1L* and *HMGA2*, which encode the master regulators of epithelial to mesenchyme transition induced by TGFβ. We thus examined the possible influence of TGFβ signaling on the re-orientation of chromatin fiber interactions. Surprisingly, after treated with TGFβ for 24 hours, both ES and EB cells lost very similar interactors as in PARG treatment. We proposed that TGFβ signaling might target PAR chains and disconnect the PAR-dependent chromatin physical association to re-organize the chromosomal interaction networks.

TGFβ signaling reduced poly(ADP-ribose) level by antagonizing CTCF-PARP1 feed-back loop

PAR has a rapid turn-over rate balanced by the activities of PARG and PAR polymerases (PARPs) in the cell. In the TGFβ-treated ES and MCF7 cells, PAR levels were reduced with different kinetics. Since the PARG activity was constant and the level of PAR precursor NAD+ increased after TGFβ treatment, we turned our attention to PARPs. We have noticed that the protein levels of both CTCF and

PARP1, and PARP activity reduced in the treated MCF7 cells. To explain the connection between the CTCF, PARP1 and PAR level, we performed *in vitro* analysis using recombinant CTCF and PARP1. The incubation of CTCF with PARP1 activated PARP1 to synthesize PAR chains from biotin-NAD+; and activated PARP1 PARylated CTCF thus generating a feed-back loop. As CTCF dissociated from PARP1 and the levels of both proteins decreased in TGFβ-treated ES cells, it was suggested that TGFβ signaling reduces PAR level by targeting the CTCF-PARP1 feed-back loop.

Summary and perspective

In this study, we have identified poly(ADP-ribose) as a molecular bridge to mediate inter-chromosomal interactions. Interestingly, TGFβ-treated cells lost these interactions by targeting the CTCF-PARP1 feed-back loop. This mechanism may be applied more broadly since many CTCF binding sites in the genome carry the PAR mark. In this sense, TGFβ signaling may rapidly rewire the overall chromosomal interaction networks by regulating PAR level. These results may form the basis for a model in which chromosomal interaction networks can be re-organized by the coordination of multiple external signaling that target or antagonize different perspectives of cell metabolism.

Paper IV

Our understanding about higher order chromatin structure and organization is currently based on the 3D DNA FISH method and 'C' technologies. The 3D DNA FISH is limited by its low resolution; while the 'C' technologies can only generate an average picture of a cell population. To bridge this conundrum, we developed a novel method Chromatin *In Situ* Proximity (ChrISP), to quantitatively visualize proximities between chromatin fibers or between chromatin fiber and structural hallmarks in single cell at a high resolution.

Development of a method to visualize chromatin fiber proximity in single cell

The principle of ChrISP is based on the visualization of two different chromatin regions within 170 Å. Briefly, probes representing two chromatin regions were labeled with biotin or digoxigenin, respectively. The labeled molecules were recognized by primary antibodies followed by secondary $F(ab')_2$ fragments conjugated with oligo sequences to generate priming and non-priming moieties. Following hybridization of fluorescently labeled splinter and backbone to bridge the moieties, a circular DNA was generated by ligation. Only when two chromatin regions are juxtaposition to each other, the circular DNA can be stabilized and fluorescent signals be visualized. The *in vitro* test on DNA fibers using a similar approach demonstrated that the method could reach a high resolution < 170 Å and that it recognizes proximity between neighboring chromatin fibers in 100% of the cells. The ChrISP method was further adapted to identify the proximity between DNA sequence and associated chromatin marks or structural hallmarks of the nucleus.

Developmentally acquired chromosome 11 large hubs impinging on nuclear membrane mediated by H3K9me2

In this study, ChrISP method was applied to analyze the higher order chromatin organization of interphase chromosome 11. In HCT116 cells, the ChrISP signals of chromosome 11 can be found almost exclusively at the nuclear periphery suggesting increased chromatin compaction near nuclear membrane. The signals – up to 0.7 micron in size – likely represent large clusters of chromatin hubs involving millions of bp of sequences. As these chromatin hubs were reminiscent of Large Organized Chromatin K9 Modification (LOCKs), which are large domains enriched for H3K9me2 predicted to be heterochromatin regions on the nuclear membrane, we performed ChrISP analysis of the spatial distribution of H3K9me2 on chromosome 11. Similar to the large chromatin hubs at the nuclear membrane, the chromosome 11-H3K9me2 ChrISP signals were also displayed as large blocks on the nuclear membrane. These chromosome 11-H3K9me2 blocks can only be visualized on the periphery but not inside the chromosome, which may due to the clustering of chromatin itself at these regions to amplify the signals. We thus proposed that the chromatin hubs on the membrane are enriched for H3K9me2 mark and represent LOCKs. Furthermore, we observed these hubs projecting to the nuclear membrane as finger-like structure providing the first visualization of a higher order chromatin conformation of LOCKs. The absence of H3K9me2 enriched chromatin hubs in human ES cells, but their emergence in EB cells provide strong visual support to the notion that LOCKs are emerging during

development to facilitate the formation of robust phenotypes.

We next addressed the possibility that the formation of such chromatin hubs is mediated by the H3K9me2 mark. When HCT116 cells were treated with either G9a/GLP methyltransferases inhibitor or G9a siRNA, the level of H3K9me2 and derived H3K9me3 were significantly decreased, as expected. Strikingly, the ChrISP signals for chromosome 11 hubs on the nuclear membrane were significantly reduced accompany with the emergence of new ChrISP signals inside the chromosome territory. We concluded that the H3K9me2 and/or H3K9me3 can mediate the compaction of chromatin regions impinging on the nuclear membrane, and the loss of such chromatin marks re-organizes chromosome-wide chromatin structures.

Anchorage of chromosome 11 to the nuclear lamina

The re-organization of chromatin structure may involve the rearrangement of anchorage to structural hallmarks such as nuclear lamina. We examined this possibility by applying ChrISP on chromosome 11-lamin A/C in control and inhibitor or siRNA treated HCT116 cells. The proximities between chromosome 11 and lamin A/C did not change with these treatments demonstrating that large-scale changes in chromosome structures do not involve detachment of chromosome 11 from the nuclear lamina. Interestingly, the lamin A/C-CT11 ChrISP signals displayed either spotty heterogeneous or more clustered patterns suggesting that the anchoring of chromosome 11 to the membrane is very dynamic.

Summary and perspective

The Chromatin *In Situ* Proximity (ChrISP) technique is a powerful new method to visualize the proximities between chromatin fibers or between chromatin fiber and structural hallmarks in single cells at a high resolution. We have used it to identify developmentally regulated chromosome 11 specific large chromatin hubs on the nuclear membrane. The hubs are enriched for and mediated by H3K9me2 and probably represent LOCKs. The removal of this chromatin mark can disassemble these hubs and re-organize the whole chromosome structure without its detachment from the nuclear lamina (Figure 10). Our study reveals that epigenetic marks regulated by developmental or environmental signaling can prevent large-scale rewiring of higher order chromatin conformations to potentially contribute to the emergence of robust phenotypes during development windows. Conversely, the unscheduled reprogramming of H3K9me2 can lead to instability in higher order chromatin structures to contribute to pathogenesis.

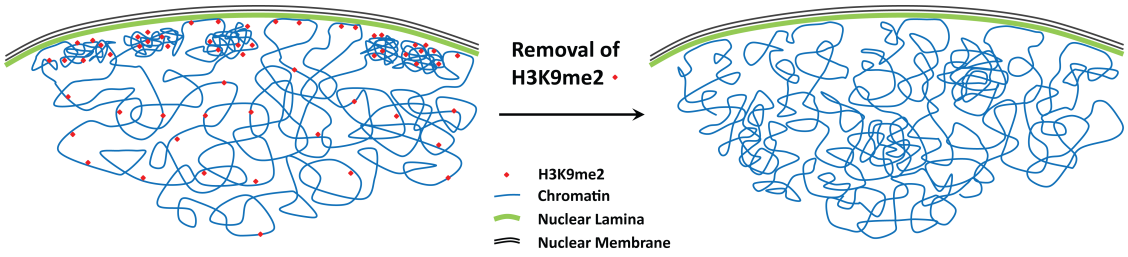

Figure 10. Model of large-scale re-organization of whole chromosome 11 structure after removal of H3K9me2. The large chromatin hubs on the nuclear membrane are disassembled without the detachment of chromosome 11 from the nuclear lamina.

4 CONCLUDING REMARKS

Using the Circular Chromosome Conformation Capture (4C) assay, we have identified global chromosomal interactomes to *H19* imprinting control region (ICR) in multiple cell types in mouse and human. The chromosomal interactome is re-organized during development, and is poorly conserved between mouse and human.

The ability to capture multiple interactions simultaneously enables us to construct chromosomal interaction networks with crosswise connections. Based on the degree of connectivity of each node, we have shown that such network in human ES and EB cells has a modular topology and is scale-free. The presentation of chromosomal interaction network topology may offer novel perspective to understand its connection to the function outcomes.

We have also identified poly(ADP-ribose) as the bridging molecule to mediate chromatin fiber physical communications. External signaling can rewire the chromosomal interactome extensively. Thus, TGFβ rewires an inter-chromosomal interactome by targeting the CTCF-PARP1 feed-back loop to reduce poly(ADP-ribose) levels.

We further captured a developmentally conserved imprinted interaction network (male germline) in mouse, which is dependent on the CTCF binding sites on maternal *H19* ICR. This physical network functions as a vehicle to transvect epigenetic states – here represented by replication timing patterns – from the *H19*/*Igf2* domain to other imprinted domains in the network. The transvection of epigenetic states *via* chromatin fiber physical interactions may trigger the spreading of imprinting and be essential for the evolution of imprinted states.

Finally, we invented a novel method, termed Chromatin *In Situ* Proximity (ChrISP), which enables the visualization of proximities between chromatin fibers or between chromatin fiber and structural hallmarks in single cells at a high resolution. Using this method, we have identified developmentally regulated large chromatin hubs impinging on the nuclear membrane mediated by the H3K9 methylation mark. Modification of epigenetic marks, which may be triggered by developmental or environmental signaling, can cause large-scale changes in higher order chromatin conformations.

Work presented in this thesis has shed new light on the understanding about the nature of higher order chromatin conformations and networks and their functional outcomes in the regulation of developmental and pathological processes.

5 ACKNOWLEDGEMENTS

The work presented in this thesis was performed at the Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institutet and at Evolutionary Biology Centre (EBC), Uppsala University. During the years of study, many people have helped and supported me. Here I want to express my gratitude towards:

My supervisor, Prof. **Rolf Ohlsson**. Over the years, you have helped and educated me a lot both on the scientific and personal level. Your enthusiasm to science, open-mind attitude and endless curiosity have greatly influenced my view of science. Thank you for all the guidance and encouragement!

My co-supervisors, Dr. **Anita Göndör**, Dr. **Mikael Sjölinder**, Prof. **Tomas Ekström** and Prof. **Laszlo Szekely**. Thank you Anita for all the scientific discussions and help in my projects! Mikael was also the supervisor for my Master thesis, thank you for teaching me various lab techniques and discussing scientific ideas! Thank you Tomas for the help on my projects, especially the *Igf2* plasmid! Thank you Laszlo for the intensive discussions about image analysis and microscopy skills, you can always get nice and beautiful pictures!

Current and previous members of the Ohlsson lab. Thank you all for the great job to keep the lab running, and the great time we spent together! **Samer**, we started our doctoral studies almost at the same time and we shared lots of great memories. Now it's your turn and good luck with your thesis! **Nori**, what about the idea that we work together and build up the FISH&ChIPs group. Thank you for the daily experimental guidance and the weekly Japanese sense of humor! **Xingqi**, you are doing a great job for both being a good father and a hard-working scientist, this shows your multi-project ability! **Moumita**, thank you for the English grammar lesson, that's not true, actually I want to thank you for the wonderful dinners and parties! **Piri**, thank you for teaching me everything about stem cells and for all the lab management work to maintain a nice working environment! **Kuljeet**, thank you for the collaboration on our projects and late night discussions when only we were in the office in Uppsala! **Farzaneh Shahin Varnoosfaderani**, OK Feri is a much easier name, stop taking photo and video of me, you have enough materials to show on my dissertation party! **Li-Sophie**, your Mandarin and Taiji are already good as native, Cantonese is better than mine. Want to learn my dialect? **Alex**, I know bioinformatics is just about pushing one button, so what are you doing everyday in front of the computer? **Anna**, still wondering how you made that birthday cake, can you have birthday every month? **Honglei**, you are still missing one journal club, please do it I like that paper. **Gozde**, 1. don't leave the keys on the door; 2. don't leave the keys on the table. **Maria**, how can you always publish single author Nature paper? **Fatemeh**, welcome to the club!

The MTC and KI people: Prof. **Ingemar Ernberg**, thank you for organizing so many wonderful international collaborations and conferences! I really enjoyed the symposia in DKFZ and Weizmann Institute. Prof. **Galina Selivanova**, thank you for being the mentor of my PhD study! Prof. **Lars-Gunnar Larsson**, thank you for the

discussion on the projects and the great performance of your band! **Björn**, thank you for the collaboration on the paper! **Hai**, **Hamid**, **Hong**, **Jie-zhi**, **Joanna**, **Junwei**, **Lifu**, **Marijke**, **Noemi**, **Pegah**, **Qin & Zimin**, **Qinzi**, **Sharon**, **Siti-Mariam**, **Xiao T**, **Xicong**, **Xun**, **Yunlong**, **Ying Y**, **Ying Z**. Thank you for the chats and being there!

The Uppsala MBB Master program classmates: **Han X** & **Ye**, I will always remember our good old days in Uppsala! **Linfeng**, please always keep at home a pair of chopsticks for me, just like what you did in Uppsala! **Yijing**, don't forget we still need to complete our plan to FIKA at every single Café in Uppsala! **Song & Na**, you should start to train your babies to become NBA players from now! **Han L**, **Rachal**, **Liang**, **Yao**, **Lin J**, **Hao T**, **Yiwen**, **Geng**, **Bin**, **Min**, **Minyan**, **Tangtang**, **Yuanyuan**, **Xiaodi**, **Xiaona**, **Qie**, **Xiaozhi**, **Yanhong**, **Hao C**. Thank you for all the memories, it was the best program I ever attended, because of you!

I also want to thank my friends outside the scientific world who make my life in Sweden wonderful! **Anders** and **Johanna**, the 'dinner group', we have had uncountable dinner together since October 2007, I enjoyed the food but most importantly your company! **Shelly** & **Jens**, I was responsible for convincing Shelly to come to Sweden, I hope you both agree that was the right move. **Wuming**, we need to increase the frequency of 'beer night, only we two' to the previous 'normal' level. **Michal**, **Punit**, thank you for all the dinners, parties, movies, circus, hiking!

My friends from CSSAU (Chinese Student and Scholar Association in Uppsala): **Xiang**, wish you everything the best from the bottom of my heart! **Mi**, we don't agree with each other, we argue all the time, we get the work done NICELY! **Meng**, thank you for being creative! **Simin**, why not trying an edit&publish job? **Jinzhi**, **Lei S**, **Hao L, Yuan T**, **Yifei**, **Xin**, **Yin**, **Xia**, **Xiao W**, **Tianlin**, **Dan**, **En**, **Yemao** and more. It was an extremely enjoyable experience to work with all of you arranging exciting activities!

Education section of Chinese Embassy in Sweden: Counselor **Ning Zhang**, Counselor **Lin Sun**, **Wei Wang**, **Hong Zhang**, **Rui Fan** and **Xudong Li**, thank you for the help during my PhD study and support to the work of CSSAU!

My dear friends from Uppsala: **Yuwei** & **Xue**, met Yuwei on my first day and Xue in the first month in Sweden. **Xi** & **Yu**, **Yanling** & **Di**, **Yang J**, thank you for the whole night talk! **Lin L**, **Peng**, **Hao** & **Grace**, **Wei W** & **Tengjiao**, **Loulou**, **Lei C** & **Hua**, **Shuang**, **Lei Y**, **Li**, **Wangshu**, **Anxue**, **Xiaohu**, **Anqi**, **Wei S**, **Shi**, **Juan**, and more friends anywhere in the world. Thank you for the friendship and support!

Dad and **Mom**, thank you for all your trust and support on me. Your endless love is always my advancing motivation!
爸爸妈妈，谢谢你们一直以来给予我的信任与支持。你们无尽的爱一直是我前进的动力！

Yan, I am so lucky to have you and your love. Meeting you is one of the most beautiful things in my life!

6 REFERENCES

Apostolou, E., and Thanos, D. (2008). Virus Infection Induces NF-kappaB-dependent interchromosomal associations mediating monoallelic IFN-beta gene expression. Cell *134*, 85-96.

Bantignies, F., Roure, V., Comet, I., Leblanc, B., Schuettengruber, B., Bonnet, J., Tixier, V., Mas, A., and Cavalli, G. (2011). Polycomb-dependent regulatory contacts between distant Hox loci in Drosophila. Cell *144*, 214-226.

Barlow, D.P. (2011). Genomic imprinting: a mammalian epigenetic discovery model. Annual review of genetics *45*, 379-403.

Bartolomei, M.S., and Ferguson-Smith, A.C. (2011). Mammalian genomic imprinting. Cold Spring Harbor perspectives in biology *3*.

Bednar, J., Horowitz, R.A., Grigoryev, S.A., Carruthers, L.M., Hansen, J.C., Koster, A.J., and Woodcock, C.L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proceedings of the National Academy of Sciences *95*, 14173-14178.

Berezney, R., Dubey, D.D., and Huberman, J.A. (2000). Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. Chromosoma *108*, 471-484.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes & development *23*, 781-783.

Bergström, R., Whitehead, J., Kurukuti, S., and Ohlsson, R. (2007). CTCF Regulates Asynchronous Replication of the Imprinted H19/Igf2 Domain. Cell Cycle *6*, 450-454.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes & development *16*, 6-21.

Bock, C., Tomazou, E.M., Brinkman, A.B., Muller, F., Simmer, F., Gu, H., Jager, N., Gnirke, A., Stunnenberg, H.G., and Meissner, A. (2010). Quantitative comparison of genome-wide DNA methylation mapping technologies. Nature biotechnology *28*, 1106-1114.

Bourc'his, D., and Bestor, T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature *431*, 96-99.

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. Science *294*, 2536-2539.

Boyle, S., Gilchrist, S., Bridger, J.M., Mahy, N.L., Ellis, J.A., Bickmore, W.A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. Human molecular genetics *10*, 211-219.

Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS biology *4*, e138.

Brianna Caddle, L., Grant, J.L., Szatkiewicz, J., van Hase, J., Shirley, B.J., Bewersdorf, J., Cremer, C., Arneodo, A., Khalil, A., and Mills, K.D. (2007). Chromosome neighborhood composition determines translocation outcomes after exposure to high-dose radiation in primary cells. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *15*, 1061-1073.

Brown, J.M., Green, J., das Neves, R.P., Wallace, H.A., Smith, A.J., Hughes, J., Gray, N., Taylor, S., Wood, W.G., Higgs, D.R., *et al.* (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. The Journal of cell biology *182*, 1083-1097.

Brown, J.M., Leach, J., Reittie, J.E., Atzberger, A., Lee-Prudhoe, J., Wood, W.G., Higgs, D.R., Iborra, F.J., and Buckle, V.J. (2006). Coregulated human globin genes are frequently in spatial proximity when active. The Journal of cell biology *172*, 177-187.

Burke, L.J., Zhang, R., Bartkuhn, M., Tiwari, V.K., Tavoosidana, G., Kurukuti, S., Weth, C., Leers, J., Galjart, N., Ohlsson, R., *et al.* (2005). CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin. The EMBO journal *24*, 3291-3300.

Cai, S., Han, H.J., and Kohwi-Shigematsu, T. (2003). Tissue-specific nuclear architecture and gene expression regulated by SATB1. Nature genetics *34*, 42-51.

Cai, S., Lee, C.C., and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nature genetics *38*, 1278-1288.

Cavalli, G., and Misteli, T. (2013). Functional implications of genome topology. Nature structural & molecular biology *20*, 290-299.

Chakalova, L., and Fraser, P. (2010). Organization of transcription. Cold Spring Harbor perspectives in biology *2*, a000729.

Chambeyron, S., and Bickmore, W.A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Genes & development *18*, 1119-1130.

Chen, E.S., Zhang, K., Nicolas, E., Cam, H.P., Zofall, M., and Grewal, S.I. (2008). Cell cycle control of centromeric repeat transcription and heterochromatin assembly. Nature *451*, 734-737.

Chien, R., Zeng, W., Kawauchi, S., Bender, M.A., Santos, R., Gregson, H.C., Schmiesing, J.A., Newkirk, D.A., Kong, X., Ball, A.R., Jr., *et al.* (2011). Cohesin mediates chromatin interactions that regulate mammalian beta-globin expression. The Journal of biological chemistry *286*, 17870-17878.

Chubb, J.R., Boyle, S., Perry, P., and Bickmore, W.A. (2002). Chromatin Motion Is Constrained by Association with Nuclear Compartments in Human Cells. Current Biology *12*, 439-445.

Cirio, M.C., Martel, J., Mann, M., Toppings, M., Bartolomei, M., Trasler, J., and Chaillet, J.R. (2008). DNA methyltransferase 1o functions during preimplantation development to preclude a profound level of epigenetic variation. Developmental biology *324*, 139-150.

Cook, P.R. (1999). The Organization of Replication and Transcription. Science *284*, 1790-1795.

Cremer, M., Hase, J.v., Volm, T., Brero, A., Kreth, G., Walter, J., Fischer, C., Solovei, I., Cremer, C., and Cremer, T. (2001). Nonrandom radial higher-order chromatin arrangements in nuclei of diploid human cells. Chromosome Research *9*, 541-567.

Cremer, M., Kupper, K., Wagler, B., Wizelman, L., von Hase, J., Weiland, Y., Kreja, L., Diebold, J., Speicher, M.R., and Cremer, T. (2003). Inheritance of gene density-

related higher order chromatin arrangements in normal and tumor cell nuclei. The Journal of cell biology *162*, 809-820.

Cremer, T., and Cremer, M. (2010). Chromosome territories. Cold Spring Harbor perspectives in biology *2*, a003889.

Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., Bickmore, W.A. (1999). Differences in the Localization and Morphology of Chromosomes in the Human Nucleus. The Journal of cell biology *145*, 1119-1131.

Cuthbert, G.L., Daujat, S., Snowden, A.W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P.D., Tempst, P., Bannister, A.J.*, et al.* (2004). Histone deimination antagonizes arginine methylation. Cell *118*, 545-553.

Cvackova, Z., Masata, M., Stanek, D., Fidlerova, H., and Raska, I. (2009). Chromatin position in human HepG2 cells: although being non-random, significantly changed in daughter cells. Journal of structural biology *165*, 107-117.

Davis, T.L., Trasler, J.M., Moss, S.B., Yang, G.J., and Bartolomei, M.S. (1999). Acquisition of the H19 methylation imprint occurs differentially on the parental alleles during spermatogenesis. Genomics *58*, 18-28.

Degner, S.C., Wong, T.P., Jankevicius, G., and Feeney, A.J. (2009). Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. J Immunol *182*, 44-48.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Deschenes, J., Bourdeau, V., White, J.H., and Mader, S. (2007). Regulation of GREB1 transcription by estrogen receptor alpha through a multipartite enhancer spread over 20 kb of upstream flanking sequences. The Journal of biological chemistry *282*, 17335-17339.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Doi, A., Park, I.H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S.*, et al.* (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nature genetics *41*, 1350-1353.

Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C.*, et al.* (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome research *16*, 1299-1309.

Duncan, I.W. (2002). Transvection effects in Drosophila. Annual review of genetics *36*, 521-556.

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M.*, et al.* (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43-49.

Eskiw, C.H., Rapp, A., Carter, D.R., and Cook, P.R. (2008). RNA polymerase II activity is located on the surface of protein-rich transcription factories. Journal of cell science *121*, 1999-2007.

Fazzari, M.J., and Greally, J.M. (2004). Epigenomics: beyond CpG islands. Nature reviews Genetics *5*, 446-455.

Federico, C., Scavo, C., Cantarella, C.D., Motta, S., Saccone, S., and Bernardi, G. (2006). Gene-rich and gene-poor chromosomal regions have different locations

in the interphase nuclei of cold-blooded vertebrates. Chromosoma *115*, 123-128.

Finlan, L.E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J.R., and Bickmore, W.A. (2008). Recruitment to the nuclear periphery can alter expression of genes in human cells. PLoS genetics *4*, e1000039.

Fitzpatrick, G.V., Soloway, P.D., and Higgins, M.J. (2002). Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. Nature genetics *32*, 426-431.

Folco, H.D., Pidoux, A.L., Urano, T., and Allshire, R.C. (2008). Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. Science *319*, 94-97.

Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H.*, et al.* (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. Nature *462*, 58-64.

Gerlich, D., Beaudouin, J., Kalbfuss, B., Daigle, N., Eils, R., and Ellenberg, J. (2003). Global Chromosome Positions Are Transmitted through Mitosis in Mammalian Cells. Cell *112*, 751-764.

Gilbert, D.M. (2001). Nuclear Position Leaves its Mark on Replication Timing. The Journal of cell biology *152*, 11F-16.

Gilbert, D.M., Takebayashi, S.I., Ryba, T., Lu, J., Pope, B.D., Wilson, K.A., and Hiratani, I. (2010). Space and time in the nucleus: developmental control of replication timing and chromosome architecture. Cold Spring Harbor symposia on quantitative biology *75*, 143-153.

Goetze, S., Mateos-Langerak, J., Gierman, H.J., de Leeuw, W., Giromus, O., Indemans, M.H., Koster, J., Ondrej, V., Versteeg, R., and van Driel, R. (2007). The three-dimensional structure of human interphase chromosomes is related to the transcriptome map. Molecular and cellular biology *27*, 4475-4487.

Göndör, A., and Ohlsson, R. (2009a). Chromosome crosstalk in three dimensions. Nature *461*, 212-217.

Göndör, A., and Ohlsson, R. (2009b). Replication timing and epigenetic reprogramming of gene expression: a two-way relationship? Nature reviews Genetics *10*, 269-276.

Grasser, F., Neusser, M., Fiegler, H., Thormeyer, T., Cremer, M., Carter, N.P., Cremer, T., and Muller, S. (2008). Replication-timing-correlated spatial chromatin arrangements in cancer and in primate interphase nuclei. Journal of cell science *121*, 1876-1886.

Gribnau, J., Hochedlinger, K., Hata, K., Li, E., and Jaenisch, R. (2003). Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. Genes & development *17*, 759-773.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W.*, et al.* (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948-951.

Guo, C., Yoon, H.S., Franklin, A., Jain, S., Ebert, A., Cheng, H.L., Hansen, E., Despo, O., Bossen, C., Vettermann, C.*, et al.* (2011). CTCF-binding elements mediate control of V(D)J recombination. Nature *477*, 424-430.

Habermann, F.A., Cremer, M., Walter, J., Kreth, G., von Hase, J., Bauer, K., Wienberg, J., Cremer, C., Cremer, T., and Solovei, I. (2001). Arrangements of macro- and microchromosomes in chicken cells. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *9*, 569-584.

Haines, T.R., Rodenhiser, D.I., and Ainsworth, P.J. (2001). Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. Developmental biology *240*, 585-598.

Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U.C., Cesari, F., Lee, C., Almouzni, G., Schneider, R., and Surani, M.A. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature *452*, 877-881.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. Mechanisms of development *117*, 15-23.

Hakim, O., Resch, W., Yamane, A., Klein, I., Kieffer-Kwon, K.R., Jankovic, M., Oliveira, T., Bothmer, A., Voss, T.C., Ansarah-Sobrinho, C.*, et al.* (2012). DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature *484*, 69-74.

Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F.*, et al.* (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. Nature genetics *43*, 630-638.

Hassa, P.O., Haenni, S.S., Elser, M., and Hottiger, M.O. (2006). Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? Microbiology and molecular biology reviews : MMBR *70*, 789-829.

Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development *129*, 1983-1993.

Heitz, E. (1928). Das Heterochromatin der Moose. I Jahrb wiss Bot *69*, 762–818.

Hepperger, C., Mannes, A., Merz, J., Peters, J., and Dietzel, S. (2008). Three-dimensional positioning of genes in mouse cell nuclei. Chromosoma *117*, 535-551.

Hewitt, S.L., High, F.A., Reiner, S.L., Fisher, A.G., and Merkenschlager, M. (2004). Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. European journal of immunology *34*, 3604-3613.

Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R., and Sasaki, H. (2008). Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. Genes & development *22*, 1607-1616.

Hiratani, I., and Gilbert, D.M. (2009). Replication timing as an epigenetic mark. Epigenetics *4*, 93-97.

Hiratani, I., Ryba, T., Itoh, M., Rathjen, J., Kulik, M., Papp, B., Fussner, E., Bazett-Jones, D.P., Plath, K., Dalton, S.*, et al.* (2010). Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. Genome research *20*, 155-169.

Hiratani, I., Ryba, T., Itoh, M., Yokochi, T., Schwaiger, M., Chang, C.W., Lyou, Y., Townes, T.M., Schubeler, D., and Gilbert, D.M. (2008). Global reorganization of

replication domains during embryonic stem cell differentiation. PLoS biology *6*, e245.
Holwerda, S., and de Laat, W. (2012). Chromatin loops, gene positioning, and gene expression. Frontiers in genetics *3*, 217.
Horsthemke, B., and Wagstaff, J. (2008). Mechanisms of imprinting of the Prader-Willi/Angelman region. American journal of medical genetics Part A *146A*, 2041-2052.
Hou, C., Dale, R., and Dean, A. (2010). Cell type specificity of chromatin organization mediated by CTCF and cohesin. Proceedings of the National Academy of Sciences of the United States of America *107*, 3651-3656.
Hou, C., Li, L., Qin, Z.S., and Corces, V.G. (2012). Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Molecular cell *48*, 471-484.
Howell, C.Y., Bestor, T.H., Ding, F., Latham, K.E., Mertineit, C., Trasler, J.M., and Chaillet, J.R. (2001). Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell *104*, 829-838.
Hu, Q., Kwon, Y.S., Nunez, E., Cardamone, M.D., Hutt, K.R., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G.*, et al.* (2008). Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. Proceedings of the National Academy of Sciences of the United States of America *105*, 19199-19204.
Hu, Y., Kireev, I., Plutz, M., Ashourian, N., and Belmont, A.S. (2009). Large-scale chromatin structure of inducible genes: transcription on a condensed, linear template. The Journal of cell biology *185*, 87-100.
Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. Journal of cell science *109 (Pt 6)*, 1427-1436.
Irizarry, R.A., Ladd-Acosta, C., Carvalho, B., Wu, H., Brandenburg, S.A., Jeddeloh, J.A., Wen, B., and Feinberg, A.P. (2008). Comprehensive high-throughput arrays for relative methylation (CHARM). Genome research *18*, 780-790.
Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature *466*, 1129-1133.
Jackson, D.A., and Cook, P.R. (1985). Transcription occurs at a nucleoskeleton. The EMBO journal *4*, 919-925.
Jackson, D.A., Hassan, A.B., Errington, R.J., and Cook, P.R. (1993). Visualization of focal sites of transcription within human nuclei. The EMBO journal *12*, 1059-1065.
Jackson, D.A., Iborra, F.J., Manders, E.M., and Cook, P.R. (1998). Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. Molecular biology of the cell *9*, 1523-1536.
Jackson, D.A., and Pombo, A. (1998). Replicon Clusters Are Stable Units of Chromosome Structure: Evidence That Nuclear Organization Contributes to the Efficient Activation and Propagation of S Phase in Human Cells. The Journal of cell biology *140*, 1285-1295.
Jeffery, L., and Nakielny, S. (2004). Components of the DNA methylation system of chromatin control are RNA-binding proteins. The Journal of biological chemistry *279*, 49479-49487.

Kaneda, A., Wang, C.J., Cheong, R., Timp, W., Onyango, P., Wen, B., Iacobuzio-Donahue, C.A., Ohlsson, R., Andraos, R., Pearson, M.A*., et al.* (2007). Enhanced sensitivity to IGF-II signaling links loss of imprinting of IGF2 to increased cell proliferation and tumor risk. Proceedings of the National Academy of Sciences of the United States of America *104*, 20926-20931.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900-903.

Killian, J.K., Byrd, J.C., Jirtle, J.V., Munday, B.L., Stoskopf, M.K., MacDonald, R.G., and Jirtle, R.L. (2000). M6P/IGF2R imprinting evolution in mammals. Molecular cell *5*, 707-716.

Kim, Y.J., Cecchini, K.R., and Kim, T.H. (2011). Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. Proceedings of the National Academy of Sciences of the United States of America *108*, 7391-7396.

Kimura, H., Tao, Y., Roeder, R.G., and Cook, P.R. (1999). Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. Molecular and cellular biology *19*, 5383-5392.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-cell dynamics of genome-nuclear lamina interactions. Cell *153*, 178-192.

Kireev, I., Lakonishok, M., Liu, W., Joshi, V.N., Powell, R., and Belmont, A.S. (2008). In vivo immunogold labeling confirms large-scale chromatin folding motifs. Nature methods *5*, 311-313.

Kitamura, E., Blow, J.J., and Tanaka, T.U. (2006). Live-cell imaging reveals replication of individual replicons in eukaryotic replication factories. Cell *125*, 1297-1308.

Kitsberg, D., Selig, S., Brandeis, M., Simon, I., Keshet, I., Driscoll, D.J., Nicholls, R.D., and Cedar, H. (1993). Allele-specific replication timing of imprinted gene regions. Nature *364*, 459-463.

Koehler, D., Zakhartchenko, V., Froenicke, L., Stone, G., Stanyon, R., Wolf, E., Cremer, T., and Brero, A. (2009). Changes of higher order chromatin arrangements during major genome activation in bovine preimplantation embryos. Experimental cell research *315*, 2053-2063.

Konig, P., Braunfeld, M.B., Sedat, J.W., and Agard, D.A. (2007). The three-dimensional structure of in vitro reconstituted Xenopus laevis chromosomes by EM tomography. Chromosoma *116*, 349-372.

Kono, T., Obata, Y., Wu, Q., Niwa, K., Ono, Y., Yamamoto, Y., Park, E.S., Seo, J.S., and Ogawa, H. (2004). Birth of parthenogenetic mice that can develop to adulthood. Nature *428*, 860-864.

Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G., and Singh, H. (2002). Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science *296*, 158-162.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929-930.

Kumaran, R.I., and Spector, D.L. (2008). A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. The Journal of cell biology *180*, 51-65.

Kurukuti, S., Tiwari, V.K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenkov, V., Reik, W., and Ohlsson, R. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. Proceedings of the National Academy of Sciences of the United States of America *103*, 10684-10689.

Leonhardt, H., Rahn, H.P., Weinzierl, P., Sporbert, A., Cremer, T., Zink, D., and Cardoso, M.C. (2000). Dynamics of DNA Replication Factories in Living Cells. The Journal of cell biology *149*, 271-280.

Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J.*, et al.* (2012). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell *148*, 84-98.

Li, J.Y., Lees-Murdock, D.J., Xu, G.L., and Walsh, C.P. (2004). Timing of establishment of paternal methylation imprints in the mouse. Genomics *84*, 952-960.

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., and Ferguson-Smith, A.C. (2008). A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Developmental cell *15*, 547-557.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Lin, C., Yang, L., Tanasa, B., Hutt, K., Ju, B.G., Ohgi, K., Zhang, J., Rose, D.W., Fu, X.D., Glass, C.K.*, et al.* (2009). Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. Cell *139*, 1069-1083.

Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qiu, X.W., Cherry, A.M., and Hoffman, A.R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science *312*, 269-272.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M.*, et al.* (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315-322.

Liu, H., Huang, J., Wang, J., Jiang, S., Bailey, A.S., Goldman, D.C., Welcker, M., Bedell, V., Slovak, M.L., Clurman, B.*, et al.* (2008). Transvection mediated by the translocated cyclin D1 locus in mantle cell lymphoma. The Journal of experimental medicine *205*, 1843-1858.

Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H., and Trasler, J.M. (2002). Methylation dynamics of imprinted genes in mouse germ cells. Genomics *79*, 530-538.

Majumder, P., and Boss, J.M. (2011). Cohesin regulates MHC class II genes through interactions with MHC class II insulators. J Immunol *187*, 4236-4244.

Majumder, P., Gomez, J.A., Chadwick, B.P., and Boss, J.M. (2008). The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. The Journal of experimental medicine *205*, 785-798.

Manders, E.M., Kimura, H., and Cook, P.R. (1999). Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. The Journal of cell biology *144*, 813-821.

Marko, M., Hsieh, C., Schalek, R., Frank, J., and Mannella, C. (2007). Focused-ion-beam thinning of frozen-hydrated biological specimens for cryo-electron microscopy. Nature methods *4*, 215-217.

Masui, O., Bonnet, I., Le Baccon, P., Brito, I., Pollex, T., Murphy, N., Hupe, P., Barillot, E., Belmont, A.S., and Heard, E. (2011). Live-cell chromosome dynamics and outcome of X chromosome pairing events during ES cell differentiation. Cell *145*, 447-458.

Mathas, S., and Misteli, T. (2009). The dangers of transcription. Cell *139*, 1047-1049.

Maya-Mendoza, A., Olivares-Chauvet, P., Shaw, A., and Jackson, D.A. (2010). S phase progression in human cells is dictated by the genetic continuity of DNA foci. PLoS genetics *6*, e1000900.

Mayer, R., Brero, A., von Hase, J., Schroeder, T., Cremer, T., and Dietzel, S. (2005). Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. BMC cell biology *6*, 44.

McBryant, S.J., Adams, V.H., and Hansen, J.C. (2006). Chromatin architectural proteins. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *14*, 39-51.

Meaburn, K.J., and Misteli, T. (2008). Locus-specific and activity-independent gene repositioning during early tumorigenesis. The Journal of cell biology *180*, 39-50.

Meister, P., Taddei, A., Ponti, A., Baldacci, G., and Gasser, S.M. (2007). Replication foci dynamics: replication patterns are modulated by S-phase checkpoint kinases in fission yeast. The EMBO journal *26*, 1315-1326.

Melnik, S., Deng, B., Papantonis, A., Baboo, S., Carr, I.M., and Cook, P.R. (2011). The proteomes of transcription factories containing RNA polymerases I, II or III. Nature methods *8*, 963-968.

Mitelman, F., Johansson, B., and Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. Nature reviews Cancer *7*, 233-245.

Moen, P.T., Jr., Johnson, C.V., Byron, M., Shopland, L.S., de la Serna, I.L., Imbalzano, A.N., and Lawrence, J.B. (2004). Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis. Molecular biology of the cell *15*, 197-206.

Moindrot, B., Audit, B., Klous, P., Baker, A., Thermes, C., de Laat, W., Bouvet, P., Mongelard, F., and Arneodo, A. (2012). 3D chromatin conformation correlates with replication timing and is conserved in resting cells. Nucleic acids research *40*, 9470-9481.

Mondal, T., Rasmussen, M., Pandey, G.K., Isaksson, A., and Kanduri, C. (2010). Characterization of the RNA content of chromatin. Genome research *20*, 899-907.

Moore, T., and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. Trends in Genetics *7*, 45-49.

Muchardt, C., Guilleme, M., Seeler, J.S., Trouche, D., Dejean, A., and Yaniv, M. (2002). Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. EMBO reports *3*, 975-981.

Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nature genetics *36*, 889-893.

Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science *322*, 1717-1720.

Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M.*, et al.* (2007). PGC7/Stella protects against DNA demethylation in early embryogenesis. Nature cell biology *9*, 64-71.

Nathan, D., Ingvarsdottir, K., Sterner, D.E., Bylebyl, G.R., Dokmanovic, M., Dorsey, J.A., Whelan, K.A., Krsmanovic, M., Lane, W.S., Meluh, P.B.*, et al.* (2006). Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. Genes & development *20*, 966-976.

Nativio, R., Wendt, K.S., Ito, Y., Huddleston, J.E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J.M., and Murrell, A. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. PLoS genetics *5*, e1000739.

Nelson, C.J., Santos-Rosa, H., and Kouzarides, T. (2006). Proline isomerization of histone H3 regulates lysine methylation and gene expression. Cell *126*, 905-916.

Nemeth, A., Guibert, S., Tiwari, V.K., Ohlsson, R., and Langst, G. (2008). Epigenetic regulation of TTF-I-mediated promoter-terminator interactions of rRNA genes. The EMBO journal *27*, 1255-1265.

Nowak, S.J., and Corces, V.G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends in genetics : TIG *20*, 214-220.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W.*, et al.* (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nature genetics *36*, 1065-1071.

Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., and Fraser, P. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS biology *5*, e192.

Paddy, M.R., Belmont, A.S., Saumweber, H., Agard, D.A., and Sedat, J.W. (1990). Interphase nuclear envelope lamins form a discontinuous network that interacts with only a fraction of the chromatin in the nuclear periphery. Cell *62*, 89-106.

Pandey, R.R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-Dinardo, D., and Kanduri, C. (2008). Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Molecular cell *32*, 232-246.

Papantonis, A., Kohro, T., Baboo, S., Larkin, J.D., Deng, B., Short, P., Tsutsumi, S., Taylor, S., Kanki, Y., Kobayashi, M.*, et al.* (2012). TNFalpha signals through

specialized factories where responsive coding and miRNA genes are transcribed. The EMBO journal *31*, 4404-4414.

Parada, L.A., McQueen, P.G., and Misteli, T. (2004). Tissue-specific spatial organization of genomes. Genome Biol *5*, R44.

Parada, L.A., McQueen, P.G., Munson, P.J., and Misteli, T. (2002). Conservation of Relative Chromosome Positioning in Normal and Cancer Cells. Current Biology *12*, 1692-1697.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T.*, et al.* (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422-433.

Pastor, W.A., Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P.*, et al.* (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature *473*, 394-397.

Pelizzola, M., and Ecker, J.R. (2011). The DNA methylome. FEBS letters *585*, 1994-2000.

Peng, J.C., and Karpen, G.H. (2007). H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. Nature cell biology *9*, 25-35.

Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W., Solovei, I., Brugman, W., Graf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M.*, et al.* (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. Molecular cell *38*, 603-613.

Pombo, A., and Cook, P.R. (1996). The localization of sites containing nascent RNA and splicing factors. Experimental cell research *229*, 201-203.

Pope, B.D., Aparicio, O.M., and Gilbert, D.M. (2013). SnapShot: Replication timing. Cell *152*, 1390-1390 e1391.

Ragoczy, T., Bender, M.A., Telling, A., Byron, R., and Groudine, M. (2006). The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes & development *20*, 1447-1457.

Rappold, G.A., Cremer, T., Hager, H.D., Davies, K.E., Müller, C.R., and Yang, T. (1984). Sex chromosome positions in human interphase nuclei as studied by in situ hybridization with chromosome specific DNA probes. Human Genetics *67*, 317-325.

Rassoulzadegan, M., Magliano, M., and Cuzin, F. (2002). Transvection effects involving DNA methylation during meiosis in the mouse. The EMBO journal *21*, 440-450.

Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature *452*, 243-247.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089-1093.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E.*, et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Rippe, K. (2001). Making contacts on a nucleic acid polymer. Trends in biochemical sciences *26*, 733-740.

Robinson, P.J., Fairall, L., Huynh, V.A., and Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. Proceedings of the National Academy of Sciences of the United States of America *103*, 6506-6511.

Roix, J.J., McQueen, P.G., Munson, P.J., Parada, L.A., and Misteli, T. (2003). Spatial proximity of translocation-prone gene loci in human lymphomas. Nature genetics *34*, 287-291.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proceedings of the National Academy of Sciences of the United States of America *105*, 8309-8314.

Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T.C., Robins, A.J., Dalton, S., and Gilbert, D.M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. Genome research *20*, 761-770.

Sadoni, N., Cardoso, M.C., Stelzer, E.H., Leonhardt, H., and Zink, D. (2004). Stable chromosomal units determine the spatial and temporal organization of DNA replication. Journal of cell science *117*, 5353-5365.

Sakatani, T., Kaneda, A., Iacobuzio-Donahue, C.A., Carter, M.G., de Boom Witzel, S., Okano, H., Ko, M.S., Ohlsson, R., Longo, D.L., and Feinberg, A.P. (2005). Loss of imprinting of Igf2 alters intestinal maturation and tumorigenesis in mice. Science *307*, 1976-1978.

Sanchez-Elsner, T., Gou, D., Kremmer, E., and Sauer, F. (2006). Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. Science *311*, 1118-1123.

Sandhu, K.S., Shi, C., Sjolinder, M., Zhao, Z., Gondor, A., Liu, L., Tiwari, V.K., Guibert, S., Emilsson, L., Imreh, M.P.*, et al.* (2009). Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. Genes & development *23*, 2598-2603.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Developmental biology *241*, 172-182.

Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. Nature *489*, 109-113.

Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature *436*, 138-141.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S.*, et al.* (2010). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nature genetics *42*, 53-61.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.

Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem *75*, 243-269.

Simon, I., Tenzen, T., Reubinoff, B.E., Hillman, D., McCarrey, J.R., and Cedar, H. (1999). Asynchronous replication of imprinted genes is established in the gametes and maintained during development. Nature *401*, 929-932.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nature genetics *38*, 1348-1354.

Solovei, I., Kreysing, M., Lanctot, C., Kosem, S., Peichl, L., Cremer, T., Guck, J., and Joffe, B. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. Cell *137*, 356-368.

Speicher, M.R., and Carter, N.P. (2005). The new cytogenetics: blurring the boundaries with molecular biology. Nature reviews Genetics *6*, 782-792.

Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. (2005). Interchromosomal associations between alternatively expressed loci. Nature *435*, 637-645.

Splinter, E., and de Laat, W. (2011). The complex transcription regulatory landscape of our genome: control in three dimensions. The EMBO journal *30*, 4345-4355.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes & development *20*, 2349-2354.

Stambrook, P.J., and Flickinger, R.A. (1970). Changes i chromosomal DNA replication tterns in developing frog embryos. The Journal of experimental zoology *174*, 101-113.

Sterner, D.E., and Berger, S.L. (2000). Acetylation of Histones and Transcription-Related Factors. Microbiology and Molecular Biology Reviews *64*, 435-459.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L.*, et al.* (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Takizawa, T., Gudla, P.R., Guo, L., Lockett, S., and Misteli, T. (2008a). Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP. Genes & development *22*, 489-498.

Takizawa, T., Meaburn, K.J., and Misteli, T. (2008b). The meaning of gene positioning. Cell *135*, 9-13.

Tan-Wong, S.M., French, J.D., Proudfoot, N.J., and Brown, M.A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. Proceedings of the National Academy of Sciences of the United States of America *105*, 5160-5165.

The 1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56-65.

The ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Thomson, I., Gilchrist, S., Bickmore, W.A., and Chubb, J.R. (2004). The Radial Positioning of Chromatin Is Not Inherited through Mitosis but Is Established De Novo in Early G1. Current Biology *14*, 166-172.

Thorvaldsen, J.L., Duran, K.L., and Bartolomei, M.S. (1998). Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. Genes & development *12*, 3693-3702.

Tiwari, V.K., Cope, L., McGarvey, K.M., Ohm, J.E., and Baylin, S.B. (2008a). A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations. Genome research *18*, 1171-1179.

Tiwari, V.K., McGarvey, K.M., Licchesi, J.D., Ohm, J.E., Herman, J.G., Schubeler, D., and Baylin, S.B. (2008b). PcG proteins, DNA methylation, and gene repression by chromatin looping. PLoS biology *6*, 2911-2927.

Tolhuis, B., Blom, M., Kerkhoven, R.M., Pagie, L., Teunissen, H., Nieuwland, M., Simonis, M., de Laat, W., van Lohuizen, M., and van Steensel, B. (2011). Interactions among Polycomb domains are guided by chromosome architecture. PLoS genetics *7*, e1001343.

Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and Interaction between Hypersensitive Sites in the Active β-globin Locus. Molecular cell *10*, 1453-1465.

Tsytsykova, A.V., Rajsbaum, R., Falvo, J.V., Ligeiro, F., Neely, S.R., and Goldfeld, A.E. (2007). Activation-dependent intrachromosomal interactions formed by the TNF gene promoter and two distal enhancers. Proceedings of the National Academy of Sciences of the United States of America *104*, 16850-16855.

Van Bortle, K., Ramos, E., Takenaka, N., Yang, J., Wahi, J.E., and Corces, V.G. (2012). Drosophila CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. Genome research *22*, 2176-2187.

van Holde, K. (1989). Chromatin. Springer-Verlag, New York.

van Holde, K., and Zlatanova, J. (1995). Chromatin Higher Order Structure: Chasing a Mirage? Journal of Biological Chemistry *270*, 8373-8376.

Waddington, C.H. (1942). Endeavour. *1*, 18-20.

Waddington, C.H. (1957). The strategy of the Genes: a Discussion of Some Aspects of Theoretical Biology. Allen and Unwin, London.

Walter, J., Schermelleh, L., Cremer, M., Tashiro, S., and Cremer, T. (2003). Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. The Journal of cell biology *160*, 685-697.

Wang, J.H., Gostissa, M., Yan, C.T., Goff, P., Hickernell, T., Hansen, E., Difilippantonio, S., Wesemann, D.R., Zarrin, A.A., Rajewsky, K.*, et al.* (2009). Mechanisms promoting translocations in editing and switching peripheral B cells. Nature *460*, 231-236.

Weksberg, R., Shuman, C., and Smith, A.C. (2005). Beckwith-Wiedemann syndrome. American journal of medical genetics Part C, Seminars in medical genetics *137C*, 12-23.

Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nature genetics *41*, 246-250.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., *et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.

Williams, R.R., Azuara, V., Perry, P., Sauer, S., Dvorkina, M., Jorgensen, H., Roix, J., McQueen, P., Misteli, T., Merkenschlager, M., *et al.* (2006). Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. Journal of cell science *119*, 132-140.

Williamson, C.M., Turner, M.D., Ball, S.T., Nottingham, W.T., Glenister, P., Fray, M., Tymowska-Lalanne, Z., Plagge, A., Powles-Glover, N., Kelsey, G., *et al.* (2006). Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster. Nature genetics *38*, 350-355.

Woodcock, C.L., and Ghosh, R.P. (2010). Chromatin higher-order structure and dynamics. Cold Spring Harbor perspectives in biology *2*, a000596.

Wu, R., Singh, P.B., and Gilbert, D.M. (2006). Uncoupling global and fine-tuning replication timing determinants for mouse pericentric heterochromatin. The Journal of cell biology *174*, 185-194.

Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F., and Barlow, D.P. (1997). Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature *389*, 745-749.

Zhang, Y., McCord, R.P., Ho, Y.J., Lajoie, B.R., Hildebrand, D.G., Simon, A.C., Becker, M.S., Alt, F.W., and Dekker, J. (2012). Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. Cell *148*, 908-921.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes & development *15*, 2343-2360.

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science *322*, 750-756.

Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., *et al.* (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nature genetics *38*, 1341-1347.

Zhu, J., Adli, M., Zou, J.Y., Verstappen, G., Coyne, M., Zhang, X., Durham, T., Miri, M., Deshpande, V., De Jager, P.L., *et al.* (2013). Genome-wide chromatin state transitions associated with developmental and environmental cues. Cell *152*, 642-654.

Zimmer, C., and Fabre, E. (2011). Principles of chromosomal organization: lessons from yeast. The Journal of cell biology *192*, 723-733.

Zink, D., Amaral, M.D., Englmann, A., Lang, S., Clarke, L.A., Rudolph, C., Alt, F., Luther, K., Braz, C., Sadoni, N., *et al.* (2004). Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. The Journal of cell biology *166*, 815-825.

7 PUBLICATIONS & MANUSCRIPTS